

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: April 10, 2004, 16:23:03 ; Search time 22322 Seconds  
(without alignments)  
17617.209 Million cell updates/sec

Title: US-09-831-000-1-copy\_60966\_70038  
Perfect score: 9073  
Sequence: 1 atggagagtcctcgatg.....ggggcaaggatcgcttga 9073

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	9073	100.0	133719	14	AF083501 Macaca mu	
2	8526.6	94.0	130733	14	AF210726 Macaca mu	
3	1714.4	18.9	1945	14	AF241163 Cercopith	
4	1607.6	17.7	137508	14	KSU75698	
5	1602.8	17.7	133661	14	U93872	
6	1002	11.0	35100	6	AR065850 Sequence	
7	1002	11.0	35100	6	AR127848 Sequence	
8	1002	11.0	35100	6	AR194750 Sequence	
9	1002	11.0	35100	6	BD190715 Unique as	
10	974.6	10.7	184427	14	EHVU20824	
11	847.4	9.3	108873	14	AF318573	
12	812.8	9.0	108409	14	AF083424	
13	770.4	8.5	112930	14	HSGEND	
14	764	8.4	113027	14	SHE410493	
15	714	7.9	35100	6	AR065851	
16	714	7.9	35100	6	AR127849	
17	714	7.9	35100	6	AR194751	
18	714	7.9	35100	6	BD190716	
19	708.4	7.8	7023	14	KSU71367	
20	679.2	7.5	60326	14	AY170316	
21	670.6	7.4	130608	14	AF005370	
22	667.6	7.4	73200	14	AF478169	
23	654.4	7.2	59673	14	AY170317	
c	24	650	7.2	2926	14	AF402655
	25	623.2	6.9	20766	14	MH68REFR
	26	623.2	6.9	119450	14	U97553
c	27	623.2	6.9	119550	14	AF105037
	28	569.4	6.3	171096	14	AY037858
	29	565	6.2	171823	14	HHV507799
c	30	565	6.2	172281	14	HSV
	31	565	6.2	184113	14	HS4958RAJ
	32	458.2	5.1	149696	14	AF319782
c	33	346.6	3.8	3445	14	AF091348
	34	346.6	3.8	3526	14	AF091349
	35	346.6	3.8	3673	14	AF091350
c	36	217	2.4	3222	14	HPU23857
	37	213.6	2.4	3140	14	HSEUNGS
	38	200.8	2.2	5829	14	HS2ULL
c	39	200.8	2.2	154746	14	HSV2HG52
	40	186.2	2.1	1772	14	AY131198
	41	184.2	2.0	30000	14	BHV130KB
c	42	184.2	2.0	135301	14	BHV130KB
	43	184.2	2.0	138390	14	AY261359
	44	183.8	2.0	110637	14	AB024414
c	45	183.8	2.0	164270	14	AB049735

ALIGNMENTS

RESULT 1  
AF083501 AF083501 133719 bp DNA linear VRL 26-JUN-2000  
LOCUS Macaca mulatta rhadinovirus 17577, complete genome.  
DEFINITION  
ACCESSION AF083501  
VERSION AF083501.3 GI:8714565  
KEYWORDS  
SOURCE Macaca mulatta rhadinovirus 17577  
ORGANISM Macaca mulatta rhadinovirus 17577  
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
Gammaherpesvirinae; Rhadinovirus.  
REFERENCE  
1 (bases 1 to 133719)  
AUTHORS Searles,R.P., Bergquam,E.P., Axthelm,M.K. and Wong,S.W.  
TITLE Sequence and genomic analysis of a Rhesus macaque rhadinovirus with similarity to Kaposi's sarcoma-associated herpesvirus/human

Pred. No. is the number of results predicted by chance to have a



herpesvirus ORF 9"		herpesvirus ORF 10"		herpesvirus ORF 11"			
/codon_start=1		/codon_start=1		/codon_start=1			
/product="DNA polymerase"		/product="DNA polymerase"		/product="DNA polymerase"			
/protein_id="AAd21336.1"		/protein_id="AAd21336.1"		/protein_id="AAd21337.1"			
/db_xref="GI:4494914"		/db_xref="GI:4494914"		/db_xref="GI:4494915"			
/translation="MDFFNPYLGRPRPHSRHTDAPAPAGAGAVQPPPPVCRLLIPA		/translation="MDFFNPYLGRPRPHSRHTDAPAPAGAGAVQPPPPVCRLLIPA		/translation="MDFFNPYLGRPRPHSRHTDAPAPAGAGAVQPPPPVCRLLIPA			
KAYDVGAGMIPTTIPFPYFENGARGDVLNANERSMWTADRPAPVPPQDQSIIF		KAYDVGAGMIPTTIPFPYFENGARGDVLNANERSMWTADRPAPVPPQDQSIIF		KAYDVGAGMIPTTIPFPYFENGARGDVLNANERSMWTADRPAPVPPQDQSIIF			
HVPDVTTHIIQQALKNAGRAAGCFSTVRYNKGILKTYDVAEHPVTEITLSSGSM		HVPDVTTHIIQQALKNAGRAAGCFSTVRYNKGILKTYDVAEHPVTEITLSSGSM		HVPDVTTHIIQQALKNAGRAAGCFSTVRYNKGILKTYDVAEHPVTEITLSSGSM			
KVPAGINVTTHIIQQALKNAGRAAGCFSTVRYNKGILKTYDVAEHPVTEITLSSGSM		KVPAGINVTTHIIQQALKNAGRAAGCFSTVRYNKGILKTYDVAEHPVTEITLSSGSM		KVPAGINVTTHIIQQALKNAGRAAGCFSTVRYNKGILKTYDVAEHPVTEITLSSGSM			
STLSDSLVACGVEFESVNDVAVRFLDHGFTTGWYSCARATPRLAARDARTALEFD		STLSDSLVACGVEFESVNDVAVRFLDHGFTTGWYSCARATPRLAARDARTALEFD		STLSDSLVACGVEFESVNDVAVRFLDHGFTTGWYSCARATPRLAARDARTALEFD			
CWNEDLSVQADSDWPPYIVAPDIECTGEAGPPCATRDGDAVIOISCVFVTRREGAP		CWNEDLSVQADSDWPPYIVAPDIECTGEAGPPCATRDGDAVIOISCVFVTRREGAP		CWNEDLSVQADSDWPPYIVAPDIECTGEAGPPCATRDGDAVIOISCVFVTRREGAP			
NPNILFSVGTCDPIPTDVLBPSPSEYDMLVSFAMIRDFVDFTLGYNINFDLPYL		NPNILFSVGTCDPIPTDVLBPSPSEYDMLVSFAMIRDFVDFTLGYNINFDLPYL		NPNILFSVGTCDPIPTDVLBPSPSEYDMLVSFAMIRDFVDFTLGYNINFDLPYL			
IIPASOVNRLNRYTKIGTSIFVHEPRGGGGFMRSSVKIKIAGIPIVDIMQVCR		IIPASOVNRLNRYTKIGTSIFVHEPRGGGGFMRSSVKIKIAGIPIVDIMQVCR		IIPASOVNRLNRYTKIGTSIFVHEPRGGGGFMRSSVKIKIAGIPIVDIMQVCR			
EKLSDYKLDVARCGKXEDVSIDIPFPRSGGGRKAVGSGVMSVLDVNDL		EKLSDYKLDVARCGKXEDVSIDIPFPRSGGGRKAVGSGVMSVLDVNDL		EKLSDYKLDVARCGKXEDVSIDIPFPRSGGGRKAVGSGVMSVLDVNDL			
KLUMTHIIEISIAKIAQARVLTDGQQLRFSLLEAAARENPILPVPTEGGGG		KLUMTHIIEISIAKIAQARVLTDGQQLRFSLLEAAARENPILPVPTEGGGG		KLUMTHIIEISIAKIAQARVLTDGQQLRFSLLEAAARENPILPVPTEGGGG			
YOGATVINIPGFYDEPVLVDVFLPSYIIQAHNLCTSTWTHGDLHLHPNLTPDDY		YOGATVINIPGFYDEPVLVDVFLPSYIIQAHNLCTSTWTHGDLHLHPNLTPDDY		YOGATVINIPGFYDEPVLVDVFLPSYIIQAHNLCTSTWTHGDLHLHPNLTPDDY			
ETPLSGGVHFVFKHRESLLRLITWLEKRAIRRTLAACDDPDSLTKTILDKQOLA		ETPLSGGVHFVFKHRESLLRLITWLEKRAIRRTLAACDDPDSLTKTILDKQOLA		ETPLSGGVHFVFKHRESLLRLITWLEKRAIRRTLAACDDPDSLTKTILDKQOLA			
IKVTCNAVGTGTVAGSILPCINIAETVTLRGKTMLEKSVYEALETEDLRLRGRE		IKVTCNAVGTGTVAGSILPCINIAETVTLRGKTMLEKSVYEALETEDLRLRGRE		IKVTCNAVGTGTVAGSILPCINIAETVTLRGKTMLEKSVYEALETEDLRLRGRE			
VTARHGPRVVGVDTSIFIAICDGVSAEAVAFCDLLAARITADLPFPIKLEAKT		VTARHGPRVVGVDTSIFIAICDGVSAEAVAFCDLLAARITADLPFPIKLEAKT		VTARHGPRVVGVDTSIFIAICDGVSAEAVAFCDLLAARITADLPFPIKLEAKT			
PKCLLLTKKRYIGLVLDKMYKMGVLDLIRKACKFVQERCBAILDLVLHPEVKARA		PKCLLLTKKRYIGLVLDKMYKMGVLDLIRKACKFVQERCBAILDLVLHPEVKARA		PKCLLLTKKRYIGLVLDKMYKMGVLDLIRKACKFVQERCBAILDLVLHPEVKARA			
RLCKRPHRYVEGLFAGFIKIVELINASYLDLNSVVPFIEQLTFSTELSRPVCYDK		RLCKRPHRYVEGLFAGFIKIVELINASYLDLNSVVPFIEQLTFSTELSRPVCYDK		RLCKRPHRYVEGLFAGFIKIVELINASYLDLNSVVPFIEQLTFSTELSRPVCYDK			
TINLPHLVYKLSAEELPOVHORIPYFVDAPGSLKSLDLAEHPDYVROHQIPVAV		TINLPHLVYKLSAEELPOVHORIPYFVDAPGSLKSLDLAEHPDYVROHQIPVAV		TINLPHLVYKLSAEELPOVHORIPYFVDAPGSLKSLDLAEHPDYVROHQIPVAV			
DLYFDKLVHGAANILQCLFGNADTTVAILYNFLNVPYKLS"		DLYFDKLVHGAANILQCLFGNADTTVAILYNFLNVPYKLS"		DLYFDKLVHGAANILQCLFGNADTTVAILYNFLNVPYKLS"			
17261..18511		17261..18511		17261..18511			
/notes="ORF 10; similar to Kaposi's sarcoma-associated		/notes="ORF 10; similar to Kaposi's sarcoma-associated		/notes="ORF 11; similar to Kaposi's sarcoma-associated			
herpesvirus ORF 10"		herpesvirus ORF 10"		herpesvirus ORF 11"			
/codon_start=1		/codon_start=1		/codon_start=1			
/product="unknown"		/product="unknown"		/product="unknown"			
/protein_id="AAd21337.1"		/protein_id="AAd21337.1"		/protein_id="AAd21337.1"			
/db_xref="GI:4494915"		/db_xref="GI:4494915"		/db_xref="GI:4494915"			
/translation="MLVNLSSVVLGDWVTFHRRFSFVNLRLQTFKRGHGYARVRL		/translation="MLVNLSSVVLGDWVTFHRRFSFVNLRLQTFKRGHGYARVRL		/translation="MLVNLSSVVLGDWVTFHRRFSFVNLRLQTFKRGHGYARVRL			
PRLSDLLHQHAFGLVTRILKELPPSPDCVALIPLDGGDDADARVAPGFVLDSSRP		PRLSDLLHQHAFGLVTRILKELPPSPDCVALIPLDGGDDADARVAPGFVLDSSRP		PRLSDLLHQHAFGLVTRILKELPPSPDCVALIPLDGGDDADARVAPGFVLDSSRP			
LTVNNAAGVHTFRCLLFLKPIDLERAVTVFNGSGARSEGTPEKPTCAESLPGGP		LTVNNAAGVHTFRCLLFLKPIDLERAVTVFNGSGARSEGTPEKPTCAESLPGGP		LTVNNAAGVHTFRCLLFLKPIDLERAVTVFNGSGARSEGTPEKPTCAESLPGGP			
LAVSGASQTSFHSFVAYPTANSVACLILRLQVRPSSDDAAHEDARLSKRYVFSN		LAVSGASQTSFHSFVAYPTANSVACLILRLQVRPSSDDAAHEDARLSKRYVFSN		LAVSGASQTSFHSFVAYPTANSVACLILRLQVRPSSDDAAHEDARLSKRYVFSN			
SGNVCKASVHTLSPBRCKAQMEIIFAPGDENAEIVLQSGFVLPFTHGGVGLVYA		SGNVCKASVHTLSPBRCKAQMEIIFAPGDENAEIVLQSGFVLPFTHGGVGLVYA		SGNVCKASVHTLSPBRCKAQMEIIFAPGDENAEIVLQSGFVLPFTHGGVGLVYA			
DAEKTQPGSSAEVRLQIFQQAAGARGLAFVTVGAPEPLFVTVYVTPALLSGCTHLL		DAEKTQPGSSAEVRLQIFQQAAGARGLAFVTVGAPEPLFVTVYVTPALLSGCTHLL		DAEKTQPGSSAEVRLQIFQQAAGARGLAFVTVGAPEPLFVTVYVTPALLSGCTHLL			
RLFPNGTPTTKRLDVLAAAPCPVVRLLSADADPRDLVASPDTGALSINAFITPVG		RLFPNGTPTTKRLDVLAAAPCPVVRLLSADADPRDLVASPDTGALSINAFITPVG		RLFPNGTPTTKRLDVLAAAPCPVVRLLSADADPRDLVASPDTGALSINAFITPVG			
FPGVSAECHVSLRDNGVHERMNH"		FPGVSAECHVSLRDNGVHERMNH"		FPGVSAECHVSLRDNGVHERMNH"			
18520..19749		18520..19749		18520..19749			
/notes="ORF 11; similar to Kaposi's sarcoma-associated		/notes="ORF 11; similar to Kaposi's sarcoma-associated		/notes="ORF 11; similar to Kaposi's sarcoma-associated			
herpesvirus ORF 11"		herpesvirus ORF 11"		herpesvirus ORF 11"			
Query Match 100.0%; Score 9073; DB 14; Length 133719;		Query Match 100.0%; Score 9073; DB 14; Length 133719;		Query Match 100.0%; Score 9073; DB 14; Length 133719;			
Best Local Similarity 100.0%; Pred. No. 0;		Best Local Similarity 100.0%; Pred. No. 0;		Best Local Similarity 100.0%; Pred. No. 0;			
Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGAGAGTTCCTCGGATGGACCAACACGTCGACCAACATCCCGGGTTCATCTTGAC	60	QY	1	ATGAGAGTTCCTCGGATGGACCAACACGTCGACCAACATCCCGGGTTCATCTTGAC	60
DB	60966	ATGAGAGTTCCTCGGATGGACCAACACGTCGACCAACATCCCGGGTTCATCTTGAC	61025	DB	60966	ATGAGAGTTCCTCGGATGGACCAACACGTCGACCAACATCCCGGGTTCATCTTGAC	61025
QY	61	ATGACGTCGATGCGCAAGTCAGGGGTGTCGTGGATCAAGTCAGTCGCTGTCAATATA	120	QY	61	ATGACGTCGATGCGCAAGTCAGGGGTGTCGTGGATCAAGTCAGTCGCTGTCAATATA	120
DB	61026	ATGACGTCGATGCGCAAGTCAGGGGTGTCGTGGATCAAGTCAGTCGCTGTCAATATA	61085	DB	61026	ATGACGTCGATGCGCAAGTCAGGGGTGTCGTGGATCAAGTCAGTCGCTGTCAATATA	61085
QY	121	ACTACGAGCCACCGGAAATGGGTGGTACGACCTGGCTTCGATCGGCTGAAGACTCC	180	QY	121	ACTACGAGCCACCGGAAATGGGTGGTACGACCTGGCTTCGATCGGCTGAAGACTCC	180
DB	61086	ACTACGAGCCACCGGAAATGGGTGGTACGACCTGGCTTCGATCGGCTGAAGACTCC	61145	DB	61086	ACTACGAGCCACCGGAAATGGGTGGTACGACCTGGCTTCGATCGGCTGAAGACTCC	61145
QY	181	GGGCGGTTCGTGGTTCACGTTTATCTAATACGGGAACCTGCTGGTGGTGGAAAGT	240	QY	181	GGGCGGTTCGTGGTTCACGTTTATCTAATACGGGAACCTGCTGGTGGTGGAAAGT	240
DB	61146	GGGCGGTTCGTGGTTCACGTTTATCTAATACGGGAACCTGCTGGTGGTGGAAAGT	61205	DB	61146	GGGCGGTTCGTGGTTCACGTTTATCTAATACGGGAACCTGCTGGTGGTGGAAAGT	61205
QY	241	ACCAGCATATCGGCGCTGTACCAAAATTTAAATCTGCTGATACGGGGCGGACCATATA	300	QY	241	ACCAGCATATCGGCGCTGTACCAAAATTTAAATCTGCTGATACGGGGCGGACCATATA	300
DB	61206	ACCAGCATATCGGCGCTGTACCAAAATTTAAATCTGCTGATACGGGGCGGACCATATA	61265	DB	61206	ACCAGCATATCGGCGCTGTACCAAAATTTAAATCTGCTGATACGGGGCGGACCATATA	61265
QY	301	GGCGCAGAACTATCGCTCGCTTAAAGACGCTTCTGCCAGATCTTCAGCGCTTTT	360	QY	301	GGCGCAGAACTATCGCTCGCTTAAAGACGCTTCTGCCAGATCTTCAGCGCTTTT	360
DB	61266	GGCGCAGAACTATCGCTCGCTTAAAGACGCTTCTGCCAGATCTTCAGCGCTTTT	61325	DB	61266	GGCGCAGAACTATCGCTCGCTTAAAGACGCTTCTGCCAGATCTTCAGCGCTTTT	61325
QY	361	GGCTTTAAGAGCGGACACATATATAGCGGTGAGAAAAGCTCATCAGACCGGAGCGTA	420	QY	361	GGCTTTAAGAGCGGACACATATATAGCGGTGAGAAAAGCTCATCAGACCGGAGCGTA	420
DB	61326	GGCTTTAAGAGCGGACACATATATAGCGGTGAGAAAAGCTCATCAGACCGGAGCGTA	61385	DB	61326	GGCTTTAAGAGCGGACACATATATAGCGGTGAGAAAAGCTCATCAGACCGGAGCGTA	61385

QY	421	TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGGTTATAGTCGACATT	480
DB	61386	TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGGTTATAGTCGACATT	61445
QY	481	ATGAAAGAGGTTATGGCGAATAAACCCTTACGATGCTACCGGATATATCAACGCGAAT	540
DB	61446	ATGAAAGAGGTTATGGCGAATAAACCCTTACGATGCTACCGGATATATCAACGCGAAT	61505
QY	541	TTTGAACCCCTCTCGAATAAGACCGGACCGGTGTTTATGGAATTCCTCAATATTATTGTAATC	600
DB	61506	TTTGAACCCCTCTCGAATAAGACCGGACCGGTGTTTATGGAATTCCTCAATATTATTGTAATC	61565
QY	601	GACGAGCGCGAAACCTCTCTTATATATCTACCTACCGCTCGTCTTCTTATCTGCTTC	660
DB	61566	GACGAGCGCGAAACCTCTCTTATATATCTACCTACCGCTCGTCTTCTTATCTGCTTC	61625
QY	661	CTAAACAGCTGGCTAAATACCCCTCTTACCGCCAGGGGGCGGTTCCTGTCATAGTAGTC	720
DB	61626	CTAAACAGCTGGCTAAATACCCCTCTTACCGCCAGGGGGCGGTTCCTGTCATAGTAGTC	61685
QY	721	GTCCGTTGCGCAACGCGAGCAACCGGTTCCAGTCACTTACCAACCGGACCGGACGCAAAAG	780
DB	61686	GTCCGTTGCGCAACGCGAGCAACCGGTTCCAGTCACTTACCAACCGGACCGGACGCAAAAG	61745
QY	781	ACGAGAGATATCGTCTGCGAGAACATCTTAACATTCATGATAGGGAAGAGTCTGTATCT	840
DB	61746	ACGAGAGATATCGTCTGCGAGAACATCTTAACATTCATGATAGGGAAGAGTCTGTATCT	61805
QY	841	GAGTACGTTACTTTGGGAGAGAACTGGGGCGTGTGTTTATAAACAACAGCGCTGCACCGAT	900
DB	61806	GAGTACGTTACTTTGGGAGAGAACTGGGGCGTGTGTTTATAAACAACAGCGCTGCACCGAT	61865
QY	901	CTGCAGTTCGGACACCTCTTAAATAATTTTATAGATATATCTTCCCATTTCTTACGAGAGTC	960
DB	61866	CTGCAGTTCGGACACCTCTTAAATAATTTTATAGATATATCTTCCCATTTCTTACGAGAGTC	61925
QY	961	ATGAGTTACGTPAGACAGATTTGTCGTTCTTAAAGTAAAGTATATGATTCCTTTTAAATATAC	1020
DB	61926	ATGAGTTACGTPAGACAGATTTGTCGTTCTTAAAGTAAAGTATATGATTCCTTTTAAATATAC	61985
QY	1021	ATTGGTGGACCCGACCTCTTTTGTTCATAGCGAGGTAAAGGCTATCTTAAACAACTA	1080
DB	61986	ATTGGTGGACCCGACCTCTTTTGTTCATAGCGAGGTAAAGGCTATCTTAAACAACTA	62045
QY	1081	CACACATGCTTAAACGCTAGGGGGCGATACCAAGGACACGAGCTCTTTTACCTGTCCCGTG	1140
DB	62046	CACACATGCTTAAACGCTAGGGGGCGATACCAAGGACACGAGCTCTTTTACCTGTCCCGTG	62105
QY	1141	GTGTGCGAGGTGTTTGTGAAGCCGTTTGAAGATPACAAAGCGGGCGCTCAACCTCACCAAC	1200
DB	62106	GTGTGCGAGGTGTTTGTGAAGCCGTTTGAAGATPACAAAGCGGGCGCTCAACCTCACCAAC	62165
QY	1201	CTCACCGTGACCGAGTGGGTGACAAAAAACCCTTTTAAAGTTAAGTAAATTTATTCGAGTTT	1260
DB	62166	CTCACCGTGACCGAGTGGGTGACAAAAAACCCTTTTAAAGTTAAGTAAATTTATTCGAGTTT	62225
QY	1261	GTGGACCAAGGACATGTTCATAGTCGCAACGGAATCCACTGAACGTTTCGACCGAGTTACC	1320
DB	62226	GTGGACCAAGGACATGTTCATAGTCGCAACGGAATCCACTGAACGTTTCGACCGAGTTACC	62285
QY	1321	TTTATACCAAGTGTGTAAACACGACCGCTGTTTAAATGGAATAAACAATAAATGC	1380
DB	62286	TTTATACCAAGTGTGTAAACACGACCGCTGTTTAAATGGAATAAACAATAAATGC	62345
QY	1381	ATATGCGGTTTCAGGGTACGTTCTTTGAGTTTCAAAAGAACTCTTAGACAGCGAACTCTTC	1440
DB	62346	ATATGCGGTTTCAGGGTACGTTCTTTGAGTTTCAAAAGAACTCTTAGACAGCGAACTCTTC	62405
QY	1441	GTGGAGACGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTAAACACGCTACTG	1500
DB	62406	GTGGAGACGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTAAACACGCTACTG	62465





Db 64626 GAGGACGAGTGTCTCTGTCGGTTTATCAGCGTCGCGCTTGTCTATAGACCTTTCTTCCCA 64685  
Qy 3721 CAGCATAAACCGCAGCTCTGAAGCTTGGCAGAAAGCGAGCTTATATATAGCTGTGAA 3780  
Db 64686 CAGCATAAACCGCAGCTCTGAAGCTTGGCAGAAAGCGAGCTTATATATAGCTGTGAA 64745  
Qy 3781 CCAATCCAGCAGAGTGTGCGGTGACCCGCGCTTCCCTCTTCCACCGTCAAGATGGT 3840  
Db 64746 CCAATCCAGCAGAGTGTGCGGTGACCCGCGCTTCCCTCTTCCACCGTCAAGATGGT 64805  
Qy 3841 GTTTAAAGCAGAACTCCCGTTTTCGCCAACAGTCTAAGCACCCGCTGAGAGGAGCGGT 3900  
Db 64806 GTTTAAAGCAGAACTCCCGTTTTCGCCAACAGTCTAAGCACCCGCTGAGAGGAGCGGT 64865  
Qy 3901 GAACCCAGGTACAGTATGCGTATCTCTTTAAATAATATTTTGGAGCTTGGAGGTATCT 3960  
Db 64866 GAACCCAGGTACAGTATGCGTATCTCTTTAAATAATATTTTGGAGCTTGGAGGTATCT 64925  
Qy 3961 GTAGTCGGAGCAACGCTAAGGCTAGTCGGGTGGCTTGACCGCGATGGTAAAGGATCTTG 4020  
Db 64926 GTAGTCGGAGCAACGCTAAGGCTAGTCGGGTGGCTTGACCGCGATGGTAAAGGATCTTG 64985  
Qy 4021 GCCAAGATCAGACCTTAAATATCTCCGGTTCGGAACAAATAGACACACACATATCT 4080  
Db 64986 GCCAAGATCAGACCTTAAATATCTCCGGTTCGGAACAAATAGACACACACATATCT 65045  
Qy 4081 ATCTATCGGTGGATAAATCAGGCTTGGTTGCTCATATCCATAACCGCTTCAGCAGAGC 4140  
Db 65046 ATCTATCGGTGGATAAATCAGGCTTGGTTGCTCATATCCATAACCGCTTCAGCAGAGC 65105  
Qy 4141 CGCGAGCTTCTGTTTAAATAATCGAGAAAGATTTAAATAATCAAGCCATGAGTCGTAAG 4200  
Db 65106 CGCGAGCTTCTGTTTAAATAATCGAGAAAGATTTAAATAATCAAGCCATGAGTCGTAAG 65165  
Qy 4201 TAGCAAGTTTGTGGAGGTTCTTACACTTCTGGTGACATTTTACTCCCAACTAT 4260  
Db 65166 TAGCAAGTTTGTGGAGGTTCTTCTAACACTTCTGGTGACATTTTACTCCCAACTAT 65225  
Qy 4261 TGTTTTAAAGCCAACTCCCAATAATCCCGCTCTGGTAAATTAATTAATTAATTAATTAATTA 4320  
Db 65226 TGTTTTAAAGCCAACTCCCAATAATCCCGCTCTGGTAAATTAATTAATTAATTAATTAATTA 65285  
Qy 4321 ACTTTTGTACTTCTGTGAGAAATCCACAGAGGTGGCGCGTATTAAGCAGTTCCGCTT 4380  
Db 65286 ACTTTTGTACTTCTGTGAGAAATCCACAGAGGTGGCGCGTATTAAGCAGTTCCGCTT 65345  
Qy 4381 TCACCTACCGTTTATGTGAAAGTCTTCAAACTCGCTTAAAACTCATCAAGACTATATTCG 4440  
Db 65346 TCACCTACCGTTTATGTGAAAGTCTTCAAACTCGCTTAAAACTCATCAAGACTATATTCG 65405  
Qy 4441 ATATAATTCACAAATGGAAGCTCTTCTCGGGACACTGTAATTTAACTTTGTAATAATG 4500  
Db 65406 ATATAATTCACAAATGGAAGCTCTTCTCGGGACACTGTAATTTAACTTTGTAATAATG 65465  
Qy 4501 CTTTATAAAACCGCTTAAGGTGAATCCATTTTACATATTTCTGTTGAGAGCGATCA 4560  
Db 65466 CTTTATAAAACCGCTTAAGGTGAATCCATTTTACATATTTCTGTTGAGAGCGATCA 65525  
Qy 4561 GCCTTATATCGCAATGTTGCTGTACACAGCCCTGGACATAAAGCTAAATTTGTAATAATG 4620  
Db 65526 GCCTTATATCGCAATGTTGCTGTACACAGCCCTGGACATAAAGCTAAATTTGTAATAATG 65585  
Qy 4621 AAGAAGCAATCTCAACCGCTCGGAAATATAAAGTCACTAAGTGTATTTTACAA 4680  
Db 65586 AAGAAGCAATCTCAACCGCTCGGAAATATAAAGTCACTAAGTGTATTTTACAA 65645  
Qy 4681 CATGGATTTACTACAAGCCGGTACACATGTAATTTAAATAAATCCACATGTAATAAT 4740  
Db 65646 CATGGATTTACTACAAGCCGGTACACATGTAATTTAAATAAATCCACATGTAATAAT 65705  
Qy 4741 AAACCTAAGGTATACATCTTCTCAATGTTTATACATAATATGAAGCTTAATCTGTACT 4800  
Db 65706 AAACCTAAGGTATACATCTTCTCAATGTTTATACATAATATGAAGCTTAATCTGTACT 65765

Qy 4801 TATATAATATGTATAGAAATTTATGCCAATACAGTCACTGTAAGTATATTTTAAAGCATAAT 4860  
Db 65766 TATATAATATGTATAGAAATTTATGCCAATACAGTCACTGTAAGTATATTTTAAAGCATAAT 65825  
Qy 4861 GCAAACTCAAAATTAATAACACAAATTCACGCTTCCGATATCCGCCAGCTGTTTAAAC 4920  
Db 65826 GCAAACTCAAAATTAATAACACAAATTCACGCTTCCGATATCCGCCAGCTGTTTAAAC 65885  
Qy 4921 CCGAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTTATTTCTCAACATATG 4980  
Db 65886 CCGAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTTATTTCTCAACATATG 65945  
Qy 4981 CAACAAATTTTCAGGTTAAACACCAATGTTTAAATAATCGCAGCATGCACTTTTATAGCTA 5040  
Db 65946 CAACAAATTTTCAGGTTAAACACCAATGTTTAAATAATCGCAGCATGCACTTTTATAGCTA 66005  
Qy 5041 AGCTCTCAAGACAAATTTCTCTTCAATGATGATCATCAACCTCGTTCGTTGTTA 5100  
Db 66006 AGCTCTCAAGACAAATTTCTCTTCAATGATGATCATCAACCTCGTTCGTTGTTA 66065  
Qy 5101 ATGTCGGAATATATAGTCCAGGATACCTTCGTCATCGCCACAAATTAATTTCCAGGGCT 5160  
Db 66066 ATGTCGGAATATATAGTCCAGGATACCTTCGTCATCGCCACAAATTAATTTCCAGGGCT 66125  
Qy 5161 TGTAAATATCATCCAGTGAATCAGCTGCTAATCTCAAGATGTTGTTTCTTAATTTTATAGC 5220  
Db 66126 TGTAAATATCATCCAGTGAATCAGCTGCTAATCTCAAGATGTTGTTTCTTAATTTTATAGC 66185  
Qy 5221 CCATCCAAATCGTGAGTGGTGACACTTGGGCAAGACCCCGTAAATTTGCGACGCTCC 5280  
Db 66186 CCATCCAAATCGTGAGTGGTGACACTTGGGCAAGACCCCGTAAATTTGCGACGCTCC 66245  
Qy 5281 ACACAAACGGAGGATGCAATTTGATCGTTACAGAGGGGACAGATATTTTATTAACA 5340  
Db 66246 ACACAAACGGAGGATGCAATTTGATCGTTACAGAGGGGACAGATATTTTATTAACA 66305  
Qy 5341 TCAGAAACCGGACCGGTCAATGACGACATATGTCATTAAAGATTGGATCGTGACACTG 5400  
Db 66306 TCAGAAACCGGACCGGTCAATGACGACATATGTCATTAAAGATTGGATCGTGACACTG 66365  
Qy 5401 GTTGAATTTGACGTTGAAAAAATGTAATTCGTGTCACATGCGTAATTTTAAACCAAC 5460  
Db 66366 GTTGAATTTGACGTTGAAAAAATGTAATTCGTGTCACATGCGTAATTTTAAACCAAC 66425  
Qy 5461 ACGAGGCGGAAATAAAACGGGTCCCGGCTAAGCGCGTGAATTCGCGGACGCGGCTC 5520  
Db 66426 ACGAGGCGGAAATAAAACGGGTCCCGGCTAAGCGCGTGAATTCGCGGACGCGGCTC 66485  
Qy 5521 ATTGCAATTTCCAGATTTGTTAAATCTGTTCCGAACATACAGTCTTAAATCAAGTATAAC 5580  
Db 66486 ATTGCAATTTCCAGATTTGTTAAATCTGTTCCGAACATACAGTCTTAAATCAAGTATAAC 66545  
Qy 5581 CACGACAAAGTTCAAAACGAGGCAACCGGCTGTTTCCACCGGACATCCCGCTCAAC 5640  
Db 66546 CACGACAAAGTTCAAAACGAGGCAACCGGCTGTTTCCACCGGACATCCCGCTCAAC 66605  
Qy 5641 AACCCGAATCTCTAGGAACGACATCAAGGTTTCAGACAAATTTCCGAAACCGAAACC 5700  
Db 66606 AACCCGAATCTCTAGGAACGACATCAAGGTTTCAGACAAATTTCCGAAACCGAAACC 66665  
Qy 5701 CTAGTTCTGAGCACTCTGTTTAAACTGCGGTTAAAGAGACGAGCGGCAACACTGC 5760  
Db 66666 CTAGTTCTGAGCACTCTGTTTAAACTGCGGTTAAAGAGACGAGCGGCAACACTGC 66725  
Qy 5761 TCGTTGTGAGGATTCAGCTTAGTGGCTGCGGCTTCTGACAGGCGGCTAGTCCGCTG 5820  
Db 66726 TCGTTGTGAGGATTCAGCTTAGTGGCTGCGGCTTCTGACAGGCGGCTAGTCCGCTG 66785  
Qy 5821 GCGGCGCTCGGCACTTTTGAAGCGGCAACCGCTTCCAGAGCGGAGGAGGCGCTCT 5880  
Db 66786 GCGGCGCTCGGCACTTTTGAAGCGGCAACCGCTTCCAGAGCGGAGGAGGCGCTCT 66845

5881 TTGAGTTCGGCCATATCTCGGCATAGAGATTGCTCGCTGATCTTTTAGTGCTATA 5940  
 5884 TTGAGTTCGGCCATATCTCGGCATAGAGATTGCTCGCTGATCTTTTAGTGCTATA 5905  
 5941 TATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGAATAAGGTCGCTCTTCT 6000  
 6006 TATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGAATAAGGTCGCTCTTCT 6065  
 6001 ATCGTTTCTCTCAATCTATAGTTACATCTTTAAGTATAACAAAGACGCTAAACACCA 6060  
 6066 ATCGTTTCTCTCAATCTATAGTTACATCTTTAAGTATAACAAAGACGCTAAACACCA 6025  
 6061 GGATGGTCGACACCTTTAAAAATCGCAAAAGCGACCGGATGGCTCTATPCGCGTTTCA 6120  
 6126 GGATGGTCGACACCTTTAAAAATCGCAAAAGCGACCGGATGGCTCTATPCGCGTTTCA 6085  
 6121 TGTCTCTCAAGAGTACGATGATCTCCAGTTAGACTCTGTCTCCCTGTAAACACC 6180  
 6086 TGTCTCTCAAGAGTACGATGATCTCCAGTTAGACTCTGTCTCCCTGTAAACACC 6145  
 6181 TTAACGGGATGGAACAGCAGCATGACACCGTGAACTTCTGTATGCTCTCTAAACCTAAC 6240  
 6146 TTAACGGGATGGAACAGCAGCATGACACCGTGAACTTCTGTATGCTCTCTAAACCTAAC 6205  
 6241 CCGAAACAGAGCTAAATACCAATGACTGTACCCCTACCCGCAACGACGCGCCGCTACTA 6300  
 6206 CCGAAACAGAGCTAAATACCAATGACTGTACCCCTACCCGCAACGACGCGCCGCTACTA 6265  
 6301 TTAGACAGGGTCAGTAACCAAGCTATCTTTTAAACCCATACGTTGGAGTTTGAAGT 6360  
 6266 TTAGACAGGGTCAGTAACCAAGCTATCTTTTAAACCCATACGTTGGAGTTTGAAGT 6325  
 6361 AAACAGCTCTGTATATTTTCAAGCGCTGTATATAATGGATACCGACGACATCAAGTA 6420  
 6326 AAACAGCTCTGTATATTTTCAAGCGCTGTATATAATGGATACCGACGACATCAAGTA 6385  
 6421 ATTAACTTTTATTTTCAAGTTCAAGATTCAGGAGGGCGTGTACAGTGGTGTGGAGCA 6480  
 6386 ATTAACTTTTATTTTCAAGTTCAAGATTCAGGAGGGCGTGTACAGTGGTGTGGAGCA 6445  
 6481 TATCTCGCCAAATGGGATGGCTGATGAATCCACACTTAGTGTCTCGATCAGTGTCTGT 6540  
 6446 TATCTCGCCAAATGGGATGGCTGATGAATCCACACTTAGTGTCTCGATCAGTGTCTGT 6505  
 6541 GTACTTTACTAGTTCGGATGAACCTAAATTTGGGTTTGGTGTAGTGTGTACCAACCAAAA 6600  
 6506 GTACTTTACTAGTTCGGATGAACCTAAATTTGGGTTTGGTGTAGTGTGTACCAACCAAAA 6565  
 6601 CTGCTGGCTCGCTTTTGAAGCTCTATGCAAGCTTTAATCGTGTCCAATAGCTTGT 6660  
 6566 CTGCTGGCTCGCTTTTGAAGCTCTATGCAAGCTTTAATCGTGTCCAATAGCTTGT 6625  
 6661 TTTAAATGCAACTCTGGTACATCTTTATGACAGTGGTCCAAACCAAGATTTAA 6720  
 6626 TTTAAATGCAACTCTGGTACATCTTTATGACAGTGGTCCAAACCAAGATTTAA 6685  
 6721 AAACACCTTAACTCGCTGTCTGGTGTCTCTGATATAACCTCTCAACCAAGAAAAA 6780  
 6726 AAACACCTTAACTCGCTGTCTGGTGTCTCTGATATAACCTCTCAACCAAGAAAAA 6745  
 6781 TTTTAAACCAAACTCGCCAGGTACTGAAACGATGTTACGGAACAGTGTGAAAGTGT 6840  
 6746 TTTTAAACCAAACTCGCCAGGTACTGAAACGATGTTACGGAACAGTGTGAAAGTGT 6805  
 6841 TATCATCTCTTCACTCCCTTTGAAGTTTGGTTTACACCGACCATGTGATAGGCA 6900  
 6806 TATCATCTCTTCACTCCCTTTGAAGTTTGGTTTACACCGACCATGTGATAGGCA 6865  
 6901 CCAGCTCAGGGGGAGTTCGGATCGTGAAGGGGTAAACAATTCGTGGCGTGGGGTACTC 6960  
 6866 CCAGCTCAGGGGGAGTTCGGATCGTGAAGGGGTAAACAATTCGTGGCGTGGGGTACTC 6925  
 6961 CAGTTTATAGTCTCCGACGAGCGTACGAGTTCTTCTGCTCAATCTTACGACGCCGC 7020

67926 CAGTTTATAGTCTCCGCGAGGACGCTACGAGTTCTTCGCTCAATCTTACGACGACGCCG 67985  
 7021 ATCCACAGGGGAGTACTGAGTTAAAGTCCCGGAATCTATCGTTCGTTAAATTTGTG 7080  
 67986 ATCCACAGGGGAGTACTGAGTTAAAGTCCCGGAATCTTATCGTTCGTTAAATTTGTG 68045  
 7081 GAGCTCTCTTTAGACAGTAAACCAATGTTGCTCCGACACTCGCGTACGATGAACGCTCTCC 7140  
 68046 GAGCTCTCTTTAGACAGTAAACCAATGTTGCTCCGACACTCGCGTACGATGAACGCTCTCC 68105  
 7141 TTGCTGTGTGTGATAGACACCTGGAGTATCTTAGACAGGTATAGGGCCGCTTCCACTCT 7200  
 68106 TTGCTGTGTGTGATAGACACCTGGAGTATCTTAGACAGGTATAGGGCCGCTTCCACTCT 68165  
 7201 AAGGTTTCCGCTTGGTAGAGCGGATTCGCGGCAATTCAGCGTCAAGTGTGTTCAATTCACC 7260  
 68166 AAGGTTTCCGCTTGGTAGAGCGGATTCGCGGCAATTCAGCGTCAAGTGTGTTCAATTCACC 68225  
 7261 GACTATCATGCGGAGTGGGATAGTGAACCTTGTAATATGATTTAAAGATGGTCTTT 7320  
 68226 GACTATCATGCGGAGTGGGATAGTGAACCTTGTAATATGATTTAAAGATGGTCTTT 68285  
 7321 CCGCTTAATGGGACTCATGTCGTTTAAAGTTCACAGGTTTAAACCACTCAGAAAG 7380  
 68286 CCGCTTAATGGGACTCATGTCGTTTAAAGTTCACAGGTTTAAACCACTCAGAAAG 68345  
 7381 GTCTGTGCGTGAACCTTTTGGGACTCTCCGAAACCGCTCAGGATAGTGTACCAAG 7440  
 68346 GTCTGTGCGTGAACCTTTTGGGACTCTCCGAAACCGCTCAGGATAGTGTACCAAG 68405  
 7441 GGTCTCTTATAAACAAGCATACGTCACAAACATGACCCAGCTGACCTTAAATAAGCT 7500  
 68406 GGTCTCTTATAAACAAGCATACGTCACAAACATGACCCAGCTGACCTTAAATAAGCT 68465  
 7501 CACAGAGAACTGTCTGATGTAATTTGAATGGAAATGTCCTCCCTGGTTCATCTCCGG 7560  
 68466 CACAGAGAACTGTCTGATGTAATTTGAATGGAAATGTCCTCCCTGGTTCATCTCCGG 68525  
 7561 GCTCATCGCGATCTGATCTTTTAACTGTTCTGCTCTACCGGGTTCCTCCGGTTAA 7620  
 68526 GCTCATCGCGATCTGATCTTTTAACTGTTCTGCTCTACCGGGTTCCTCCGGTTAA 68585  
 7621 AACCCGCGGGCCCGCCACCTGTATAACGTCGCGGAGAAATGCGCCAGGCACTGTGAG 7680  
 68586 AACCCGCGGGCCCGCCACCTGTATAACGTCGCGGAGAAATGCGCCAGGCACTGTGAG 68645  
 7681 AATTTCTACCGAGAGGGCCCTTTGTTGACCGAAAGATGTTATAGCTGTGCTGCTGAG 7740  
 68646 AATTTCTACCGAGAGGGCCCTTTGTTGACCGAAAGATGTTATAGCTGTGCTGCTGAG 68705  
 7741 CGGGTCTGTTATACCGCCCGCAGTTGGCCAGAGTTCTGCACGACGTCTACCGGAAATGAA 7800  
 68706 CGGGTCTGTTATACCGCCCGCAGTTGGCCAGAGTTCTGCACGACGTCTACCGGAAATGAA 68765  
 7801 GGCCAAAGTGTGGGGCGGTGCTCTCATATGTTGAGGGCGCTATCATGGCTAT 7860  
 68766 GGCCAAAGTGTGGGGCGGTGCTCTCATATGTTGAGGGCGCTATCATGGCTAT 68825  
 7861 TGCCGACTCGCTCTGCTCACTTATATACCTCGATGCGGAGGAAACTGGAACCTTAG 7920  
 68826 TGCCGACTCGCTCTGCTCACTTATATACCTCGATGCGGAGGAAACTGGAACCTTAG 68885  
 7921 ACTCAAGCACTGTGCAAACTTTGTTTCAAACCATCTTTCTCAAGAAATCTTAGCCCC 7980  
 68886 ACTCAAGCACTGTGCAAACTTTGTTTCAAACCATCTTTCTCAAGAAATCTTAGCCCC 68945  
 7981 TATCAGCTACTGCGCAACCGGAAGATGGTTCCTGACAAATTTATTTACCATCAGCGTAC 8040  
 68946 TATCAGCTACTGCGCAACCGGAAGATGGTTCCTGACAAATTTATTTACCATCAGCGTAC 69005  
 8041 GGCCGAGAGAGGGCCCTCTGCTGACTGGAAGTACTAGCGGATGAGCTGTCCGGAGAG 8100

69006 GGCCGAGAGAGCGCCCTGTCGTGACTGGAAGTACTAGCGGATGACGCTGTCGGGAG 69065  
QY 8101 CAGCCTTGCCCGATTCCTTAATCTTCAGTATCGGAACCGGGCTTCTCCGGCACC 8160  
Db 69066 CAGCCTTGCCCGATTCCTTAATCTTCAGTATCGGAACCGGGCTTCTCCGGCACC 69125  
QY 8161 CTTGGTTGACCTCAGTAATGTTCTTAGAAATCAGAAATCATCCTCAGCGCCGCCACCT 8220  
Db 69126 CTTGGTTGACCTCAGTAATGTTCTTAGAAATCAGAAATCATCCTCAGCGCCGCCACCT 69185  
QY 8221 GAGTCATTTGTCATCACAACACGACCGCCAGTTCCTCAGTACAGTACAGTATTATAC 8280  
Db 69186 GAGTCATTTGTCATCACAACACGACCGCCAGTTCCTCAGTACAGTACAGTATTATAC 69245  
QY 8281 GCCAACCCAGGGGTTGTTCCCGCCCAATGTTTATGAGACAGTGAAGACGGTGTACA 8340  
Db 69246 GCCAACCCAGGGGTTGTTCCCGCCCAATGTTTATGAGACAGTGAAGACGGTGTACA 69305  
QY 8341 GAGCATTCACACAGGACAGACGGCTATTTGGCCGCGCACAACCGGTTCCGACATC 8400  
Db 69306 GAGCATTCACACAGGACAGACGGCTATTTGGCCGCGCACAACCGGTTCCGACATC 69365  
QY 8401 TGGGGCCCTGGCCGCATATCGCATGTTCCCGAGTTCGGGACAGTCTCGCAGGTGA 8460  
Db 69366 TGGGGCCCTGGCCGCATATCGCATGTTCCCGAGTTCGGGACAGTCTCGCAGGTGA 69425  
QY 8461 AGGCTCCGGGGCTCGATTCCGGGAAACAGCATCGCTTCGGCAGCCCGCTTACCGCC 8520  
Db 69426 AGGCTCCGGGGCTCGATTCCGGGAAACAGCATCGCTTCGGCAGCCCGCTTACCGCC 69485  
QY 8521 CAATGTCGCCGCAAGAGATGAAACGGTAGACACAGTTGGGAAACGCTCCGTAAGAAA 8580  
Db 69486 CAATGTCGCCGCAAGAGATGAAACGGTAGACACAGTTGGGAAACGCTCCGTAAGAAA 69545  
QY 8581 CGTCCATCCGAGGCGCGTATACGCTCCTCAGTGTATATATACCAATAGACTTAC 8640  
Db 69546 CGTCCATCCGAGGCGCGTATACGCTCCTCAGTGTATATATACCAATAGACTTAC 69605  
QY 8641 GTCCCCGTCAAGGTCCGGCCAGAGTCCGGCCGATATCGCAACACTCCAGATCCCGCAT 8700  
Db 69606 GTCCCCGTCAAGGTCCGGCCAGAGTCCGGCCGATATCGCAACACTCCAGATCCCGCAT 69665  
QY 8701 GGCGGCGGCTCTCCGCCCTTCGCCGAAACCGCCGCAACGGTCCCGCTTAAGAGAAGA 8760  
Db 69666 GGCGGCGGCTCTCCGCCCTTCGCCGAAACCGCCGCAACGGTCCCGCTTAAGAGAAGA 69725  
QY 8761 GCCACGCGAGAGCTGGCAGACAAAGACTGAAGGCGCAGCTTCGGGGCGCCGCAACAGT 8820  
Db 69726 GCCACGCGAGAGCTGGCAGACAAAGACTGAAGGCGCAGCTTCGGGGCGCCGCAACAGT 69785  
QY 8821 AAACCCACCTTTCGGGGACCGTCCGGGATGCGCGTTCGGGAGCAGGGCTTATTCGATT 8880  
Db 69786 AAACCCACCTTTCGGGGACCGTCCGGGATGCGCGTTCGGGAGCAGGGCTTATTCGATT 69845  
QY 8881 AATCGAAAGCTCCACGATGTAACCGCGACGATCTGACACCCCAAAACGACGACAT 8940  
Db 69846 AATCGAAAGCTCCACGATGTAACCGCGACGATCTGACACCCCAAAACGACGACAT 69905  
QY 8941 GCTAGCGGCTATCCTACAGGACCTGTATGAGCTACAGTCCCGCCCGCCATCGATCCCC 9000  
Db 69906 GCTAGCGGCTATCCTACAGGACCTGTATGAGCTACAGTCCCGCCCGCCATCGATCCCC 69965  
QY 9001 CTCAGCAACTCGGACATGAGAGATATTTCCAGAGGTTAGTCCGCGCATCTAGCGGCA 9060  
Db 69966 CTCAGCAACTCGGACATGAGAGATATTTCCAGAGGTTAGTCCGCGCATCTAGCGGCA 70025  
QY 9061 CGGATCGCCTTGA 9073  
Db 70026 CGGATCGCCTTGA 70038

RESULT 2  
AF210726

LOCUS AF210726 130733 bp DNA linear VRL 27-MAR-2000  
DEFINITION Macaca mulatta rhadinovirus 26-95 long unique region L-DNA,  
complete sequence.  
ACCESSION AF210726  
VERSION AF210726.1 GI:7329990  
KEYWORDS  
SOURCE  
ORGANISM  
Macaca mulatta rhadinovirus 26-95  
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
Gammaherpesvirinae; Rhadinovirus.  
REFERENCE  
1. (bases 1 to 130733)  
Alexander, L., Denekamp, L., Knapp, A., Auerbach, M. R., Damania, B. and  
Desrosiers, R. C.  
The primary sequence of rhesus monkey rhadinovirus isolate 26-95:  
sequence similarities to Kaposi's sarcoma-associated herpesvirus  
and rhesus monkey rhadinovirus isolate 17577  
J. Virol. 74 (7), 3388-3398 (2000)  
20173730  
MEDLINE  
PUBMED  
10708456  
REFERENCE  
2. (bases 1 to 130733)  
Alexander, L., Denekamp, L. M., Knapp, A., Auerbach, M., Czajak, S.,  
Damania, B. and Desrosiers, R. C.  
Direct Submission  
Submitted (02-DEC-1999) Microbiology, New England Regional Primate  
Research Center, One Pinehill Dr, Southborough, MA 01772, USA  
Location/Qualifiers  
1. 130733  
/organism="Macaca mulatta rhadinovirus 26-95"  
/mol\_type="genomic DNA"  
/isolate="Macaca mulatta rhadinovirus isolate 26-95"  
/db\_xref="taxon:119193"  
513..1784  
/gene="R1"  
513..1784  
/gene="R1"  
/codon\_start=1  
/product="ORF01"  
/protein\_id="AAF59980.1"  
/db\_xref="GI:7329991"  
translation="MFVLVFLMLQPVSVELLPAKLTSVPTWCPHPGSDTVLLTCRGT  
STARDQSTQWRNTLMRGSNFYGLVSVTPNATISDRYACQTKTKTNNIDFRVS  
LRTLTQERCGSSYGYTYANNTVRVLCRSGGNVTVNVLNVTGNTGRTNTHFV  
LSRTGTGFCSAFIGNEKVFQSVOTINVFTSFPTKPTNDIPNESHFNKTIQOTASV  
QHPENVYFVFPVPSIGVLGIALILMCLMFLTRCNSESTNSVASOTSYIQPSH  
NQRNTECRSHYRNAHQBSIELEFNQHTSETDSCQLVLLKVNVAVDGQENIT  
NEVMEQDVVVVENIETQSVEDNVEHMDYSDTINPNFYISGLILEEVEVFNLELN  
QYHLILENDHNEYHNLNMLEQYDWLE"  
complement (1852..2418)  
/gene="ORF02"  
complement (1852..2418)  
/gene="ORF02"  
/codon\_start=1  
/product="DHFR"  
/protein\_id="AAF59981.1"  
/db\_xref="GI:7329992"  
translation="MDIACVIVADVQLGKGNTPWVPLRNMVYFQMTSTPSV  
VGEKNVIMGRKTFWFSIPEKRPIAVNINILSRELPEPHGAHFLATLDDAFNFR  
QYKLKEQLNTVWIGKSVYSLNXYKPLKLYITRIMESFDCDVFPSINFTEYTM  
SELFPGKDTNFENGIGIKYKQVIEKNFK"  
2595..3782  
/gene="ORF04"  
2595..3782  
/gene="ORF04"  
/codon\_start=1  
/product="complement binding protein"  
/protein\_id="AAF59982.1"  
/db\_xref="GI:7329993"  
translation="MFWIPITLCFCQVILVDSKDGENVCLKPPFPRFMVKTAQNE  
NYAVTRVELTCRPFYQLQAVYVECLSNGTWTPNAECRKRKCSPEIDILNGEVII  
TDSNAPFGSNITKYKNTGYLLGATVTRCLLYKDSNLVDWQPAAPTCIEKCKQP  
DIENKXYPVQEFYNYLETITFTCNKDFSLIGNTTTCMTNGTWSSPVPKQOITCSA  
PNIDHGTLLVGSSEVYKHGOSVTIGCEDGTLNGHKMCTCEYSLWNPLPTCPINKT  
DPTSEVPSPTCKEQEMPTENPKSHSESTTTTPTKTHKSTSPSPKVPNPETHK

```

TTPEAGISKQTTNNRPSKAPSQNPPIAPPMSKMKQBHVVLVLFASVASLLFLVLAALYCC
FLK"
4213. .7611
/gene="ORF06"
CDS
4213. .7611
/gene="ORF06"
/codon_start=1
/product="esBdp"
/protein_id="AAF59983.1"
/db_xref="GI:7329994"
/translation="MASKGNAGQPLEDNQSRAPIGACGVVYAYSKQDPFFPARASILG
NRPGSVFLPLILYGLTVHEHPEPLTKAAKYKVDVTTLAVKTCFHREVIVFHNASL
FRPYDTGNELCEEARALFGYQTTEPGFPVLSINPLFCPLQPKDMEFLGVVYTE
GFKERLNRGCLVAFQVOTQVQIAGROAFKVPVLYDEDLFAPHGRMPRFHKDVSYAL
VDSLFTIAQALRLKDVATVHATEKQFMDHYKIAKIYQAKQFSTLTKPTDGSJHM
YDSVVAELALSVCMELECPQDACLINLSWPIFDGCDSPPEARVALERWSAEOAV
HVAQQLFAANSVLILTKVQQAARGQGDVNVNSFLQGLGLFLENAIKENGSAF
KGVSNALDSSFFPYSLAYAAFSPLHLAKCYMQLQHHKSSNTQAFNMVHYVGT
AANSEMCTLCHGNTFATCLANTLYRLKDRFPATVTPQRRDPYVYVTTAGTAFNDLEILG
FRKGRDEEDGNADPEPKYTYWQLCQTVTEKLSAIGITEDHNMVFNITNQSFLR
VFKGIDISIGDEVNKFVNSMKNNFNREHVKSHHILQFCNVYQWQACAVELNLYY
KSLWIIQDICLPCMYEQDNPMANGILPSEMLKMHFOTLWTFKACILDRILGTCE
LKVHREDFCDPFDTDAGSLMAPKQWVRIARANMVVPKSIKIKNRIIPSNATGSE
AVQSGFKPQGTDRITYVAGPYMKFLNSLRHALFPDQTKTAAILYMHKISQNTKTPVK
DVDEDAELVSYVYKTSNLAPEETNVDVVDPSLMSYARIKMGAILRACGQIQFAT
TLCLTPVLQITDAEYPHVLGSAAIATPAYLAELNGRTALTQVTTARQPVAAOTRL
RPVITVPMVNNKTYGVNGNHNFCNGLNYFAGRGVDRLNPSSPFKKTGVSAMLRK
RHWMTTIIIDRLIKRAAQOTISTPEARSVKRSQALIEDXDNFNLLKSVILEIRHLG
KGCODLASEDVQYVLDGYCMLTDEVLFTLQNIAGSGVPTWIEDAGALIEDQDADDLQ
FVDSDDIATASQCPPEQQLTPPSAGALLAGKRNKINALSLDL"
7636. .9696
/gene="ORF07"
7636. .9696
/gene="ORF07"
/codon_start=1
/product="transport protein"
/protein_id="AAF59984.1"
/db_xref="GI:7329995"
/translation="MARLEALYAQLSALAVDLSLVIFADPRSIDGAKILKTKTQIEN
LNRLDPLLRQNSVETSSUSLEVHELAKNITDKIGELERSLQRYSRSHPEFTELSR
PECHYRSTPYFGGGLIDVNMCLINDVELLCKGSLGVSFYICANEALSGDLRVLT
LSTLRGSPFHPHYLTVSVPCVQCLREIELVFNQGSLLIADLACHDLCKKRAAE
PIHGLFTESQLGLKVTQSDATQHGVRSSADOLRESSLAQIDNINIFKRVASIME
LSNLIYNNAGTQGLQGTQENCSOMARLLTHEADMEHRAHITPKLSATDFYCFDPD
PIESLFCGLFNSIDDTINALSRDCSVTFQQAQYNNVWKQNELFTRLSLIHQGSA
GSQAPAPSPERTTVAATAASDVIKDAQYKQYMKKVARDQFKKLTECLQIQGAVL
ANALCMRWGVSAYGEASELNVNHFLLRRFRVAFPEAKRSDQLIFENSKYIKNSLYS
QRLGREHVEIITLQFYGLITGLPTQSLDFPGPANVYLAQFEAGMLPHMKMLVSEM
IWPQIQKDWIDQTFNRFYQLPEGLDNVQKSAWCFIRELVLSVLYNRTWEXTLRIF
SLAREKLSISNLVDKGLTSLGLTYEQDAPLVLITSQNTGIFKDLVALLYVHHLQLSDG
HDDN"
9683. .12172
/gene="ORF08"
9683. .12172
/gene="ORF08"
/codon_start=1
/product="ORF08"
/protein_id="AAF59985.1"
/db_xref="GI:7329996"
/translation="MMIINTRRLLRAWVILAIQTAIVENVTTPKGAITTAKTPTGP
STTPPENPPRAEAFKFRVCASATGELFRNLNLEKTCPTGEDKTHQEGILMVPKKNIV
PHIFKRYRIVKATSVTVYRGWTEATVTKGEVIRPVQVEINHMDTYQCFSSMRVN
VNGIVNTYDRDFTNQTVLQFVEGLTDNIQRFESQVLYXTTPGWFEGYIRVRYTVC
EIVDMIASAPYSYVFVATLGDIVVEGVPFCHNDSTCSVAEKTEGLGVALTNYTVN
PATRKPTTFVFAVDSGRTVSWKAEDPKSAVCALTLWTKFPRAIQTTHASVHYFVN
DVATFTSPLSSEVANFTGYSCLDEVIOKTLNDITIKKLSDTHVNTGSAQYKTEGGLF
LWLOPFLSLVDENRGLNGTTPAPATTSVSVRSVNTNQAQNTLAAPOLOFAY
DKIRASINKVLEESACRWCEQVDRDTNMVLSKINFTSVMTALYGRPVSAKVFQDAI
SVQPCVAVDQASVSIHSLRSTPGMCSYKPPVTFPFLNSTLTKFQGLQPRNRIILTD
NVOVACETCEHYFIASNVTYYKYDYVFKKINTSEISTLGTFLALNSPIENIDRFV
IELYSRAEKLKSGSVFDIETFREYNTYQRLAGLREDLDNTIDNLDRILDRLSIV
ADLGDVGRVNVNASSVITLFGSIVSGFINFIKPFQGMILVIVAVILVIFALNR
TNAIAQAIFIRMIYPDIDKQPSGGKVDQEQIKNITLACGWHQLQOEERRRLDEQORSAPS

```

```

LFRASDGLKRRFRGKYLENEBAQEYEMSK"
12290. .15334
/gene="ORF09"
12290. .15334
/gene="ORF09"
/codon_start=1
/product="polymerase"
/protein_id="AAF59986.1"
/db_xref="GI:7329997"
/translation="MDFENPYLGRGPEPHSHRGTTDAPAPAGAVQPPDPVCLLIPA
CLYDTPGAGMIPVITPPPTTFENGARGDVLILANERSWMTARDKRPVAPDQDQSIITF
HAYDVTVAADCAEAVPSRFQDIIIPSGVILKLLGTEDEGTSVCNVNFQQVTFIYA
RYPAGINTHILQALKNTAGRAACGFRSTVNRKRIILKTYDAVHEHPVTEITLSSGML
STLSRLVACQCFEVSNDVAVRPFVLDHGTFTGWYSCARATPLAARDATLAEFD
CSWEDLSQADRSQDPPIRYIVAFDIECTGEAGFCATRDGDVAIQISCVFTYTBREGAP
NPNELRSVGTCDPDPDVLVEFPSEYDMLVSFTFAMIRDFEDFTGTGNIENFDLPYL
ITRASOVNLRINEXYKIKTQSIGIEVHEPRGGGSGFMSVSKIAGIVPIDMVOCVR
EKLSDYKALTVAQCLGKKEDEVSYKDIPPLFRSGPGGRAKVSGYCMDSVLMDL
LKMFMHVEIIEIALAKIQARRVLTDGQRLRVFSCLEAAARENFILPVTTPGCGG
YOGATVINPIPGFVDEPVLVDFASLYPSIIQAHNLCTYTMHGDHLHLNLTDPDDY
TEVLSGGGVHFVKKHRESLLGLLTLWLEKRAIRRETLAACDDPSLTKLIDKQOLA
IKVTCNAVYGTGASGLLPCINTAETVTLRGRTMLEMSKSYVEALTTEDLRLGRE
VTARHAGFRVYVGTDSLFIACDGSABAVSAFCDDLAARITADLPPIPKLEAKT
FKCLILLPKBYIGVLLNDKVMKGVDLIRKTAACKFVOCRAILLDLDHDPPEVCAAA
RLLCKRPHAVYESGLPAGFIKIVEVLNASYLDLNSVPIEQTLFSTELSPVCDDYK
TTLNPLHAYIQRKASRCELPQVHDRIPIYFVDAPGSKSLASLAEHPDIVRQHQIPVAV
DLYFDKLVHGAGNITLQCLFGNNADTTVAILNFLNVPYKLFPS"
15429. .16679
/gene="ORF10"
15429. .16679
/gene="ORF10"
/codon_start=1
/product="ORF10"
/protein_id="AAF59987.1"
/db_xref="GI:7329998"
/translation="MLVNELSVLGDWEVTFHRGRFSFVNLTRLQTFKHGGYARVRL
PESLDQLLOHFAFGLVTRKLELPPFSDCVALIAPLDGSGDADARVAPGVLDSSRP
LTVWNASGRHTIRFCLLFLKPIDLERAVTVFVGENGARGSEGTPKPTCATESLPGGP
LEVSASQSTPHSPVAYEPANSVACLSLRLQVRPFSDDAARHDAIRISPKYVTFSN
SGCNVCKASVITLSPSRCKTAQMEIILYAPGPNAEIILVQSGPVLPTHTGGRVLGVYA
DAEKTIQGSSABEVQLIFQOAAARQDGLAFVTVGAPEPLFVVTPLLALLSGCTHLL
RLFNPNGTPTTKRDTLVAAPCPFVWLSADDAADAPDLVASFDPDTGALUSINAFITPVG
PFGVVSASCHYSLRDNGVHERMNH"
16688. .17917
/gene="ORF11"
16688. .17917
/gene="ORF11"
/codon_start=1
/product="ORF11"
/protein_id="AAF59988.1"
/db_xref="GI:7329999"
/translation="MGTPVPRFPGEWOTSSLVNDGTPRYSSLVMAATIHGDYITLVNR
SELCVTERSPCLPACPSIGRLVGKFPGFASATLGDRTGTETVYAFCHRONPLDII
PAVVERADRELVRVHAQTTRVSRYGLKVFAIVTVVRPPGVFLHFPQDRVPIALTD

```

Query Match 94.0%; Score 8526.6; DB 14; Length 130733;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 8786; Conservative 0; Mismatches 259; Indels 34; Gaps 7;

QY	1	ATCGAGAGTTCGTCGGATGGACCAACACACGTCGAACCAAAATCCCGGGTTCATCTTGAAC	60
Db	58974	ATCGAGAGTTCGTCGGATGGACCAACACACGTCGAACCAAAATCCCGGGTTCATCTTGAAC	59033
QY	61	ATGACGTCGCATGCCAAAGTCAGGGGTGTGCTGGATCAGTCAGTCGCCTGTCAAATATA	120
Db	59034	ATGACGTCGCATGCCAAAGTCAGGGGTGTGCTGGATCAGTCAGTCGCCTGTCAAATATA	59093
QY	121	ACTACACAGCCACCGGAATGGTTGGTACGACCTGGGCTTCGATCCGGCTCAAGACTCC	180
Db	59094	ACTACACAGCCACCGGAATGGTTGGTACGACCTGGGCTTCGATCCGGCTCAAGACTCC	59153
QY	181	GGGCGGTTCTTCGCGTTTACCGTTTATCTPAATACGGGAACGTGCTGCTCGGAAAAGT	240



2401	Qy	ATTGTGTTTATTATTCAGTCCAAACACGCGCCACGGATAGTTGTCATTTCCACACACCCGGGG	2460
61374	Db	ATTGTGTTTATTATTCAGTCCAAACACGCGCCACGGATAGTTGTCATTTCCACACACCCGGGG	61433
2461	Qy	GGGTGGCCGCCATAGTTTCAGACACAGAAACGCCCGGCCCTCTCGGTTAGAAGTAGGCC	2520
61434	Db	GGGTGGCCGCCATAGTTTCAGACACAGAAACGCCCGGCCCTCTCGGTTAGAAGTAGGCC	61493
2521	Qy	CGCCCCCGTGGTCGAGGTCAATACGTCCTCTTAGGCCCTCTGGTTGTACAACAATAATCA	2580
61494	Db	CGCCCCCGTGGTCGAGGTCAATACGTCCTCTTAGGCCCTCTGGTTGTACAACAATAATCA	61553
2581	Qy	CGTGCATCACCGTTGACGTCCCTCGCGTACGCGGCTCTTGACAGCGGCACGAACAC	2640
61554	Db	CGTGCATCACCGTTGACGTCCCTCGCGTACGCGGCTCTTGACAGCGGCACGAACAC	61613
2641	Qy	CGCGCTTTGAACGTGATCTGCATCGTCGTCTGAAGATGAACACACAGGTCAATCA	2700
61614	Db	CGCGCTTTGAACGTGATCTGCATCGTCGTCTGAAGATGAACACACAGGTCAATCA	61673
2701	Qy	CATCAGTGGCGGCCGTGCCCTCATCGGCGTCTGATTCGTCCCATCTCTCAACAGAGCCTC	2760
61674	Db	CATCAGTGGCGGCCGTGCCCTCATCGGCGTCTGATTCGTCCCATCTCTCAACAGAGCCTC	61733
2761	Qy	GGTCGTCACATTCGTCGTCATCTGAAATAAATAGTCTCTTGTCTCTGCAGTTCATCCAG	2820
61734	Db	GGTCGTCACATTCGTCGTCATCTGAAATAAATAGTCTCTTGTCTCTGCATTTGATCCAG	61793
2821	Qy	CAACAAACAGG---TTTCTTCTTCAATCGGAAGACATCCGCGCTCTCCCGGTGATCTAGATA	2877
61794	Db	CAACAAACAGGTTCTTCTTCTTCAATCGGAAGACATCCGCGCTCTCCCGGTGATCTAGATA	61853
2878	Qy	TATCACTGACACAGGCGGCGCGCGTGGCGAAAGCCGTCAGCTTCTCCGCCACTTT	2937
61854	Db	TATCACTGACACAGGCGGCGGAGGCCGTTGGGAAAGCCGTCAGCTTCTCCGCCACTTT	61913
2938	Qy	CATCAGTGTCTCGTAGTCACTCTATAAAAAATCGTCTTTCGCTGCTGATTCAGATGACT	2997
61914	Db	CATCAGTGTCTCGTAGTCACTCTATAAAAAATCGTCTTTCGCTGCTGATTCAGATGACT	61973
2998	Qy	GGGGTGGCGGGTTCGGCAGTAACCAACGCAACGGCCGCGGACCCCAACACGGGTCA	3057
61974	Db	GGGGTGGCGGGTTCGGCAGTAACCAACGCAACGGCCGCGGACCCCAACACGGGTCA	62033
3058	Qy	GTGGCGGGCGCGCTCGGTGTGTGGTGCCACCTGTGGAAGTAATGACGTAGTAGCC	3117
62034	Db	GTGGCGGGCGCGCTCGGTGTGTGGTGCCACCTGTGGAAGTAATGACGTAGTAGCC	62093
3118	Qy	GGCGGCTCTGGTTAAAGACTACACGTCGCGTGCCTGCGGATCTATTCTAGTGGGGAGG	3177
62094	Db	GGCGGCTCTGGTTAAAGACTACACGTCGCGTGCCTGCGGATCTATTCTAGTGGGGAGG	62153
3178	Qy	AGACCGGGCGGTTCGTCTTCTTCATCCACATCACTGTAGTCGCTGTTCTCCGAATGAGAGT	3237
62154	Db	AGACCGGGCGGTTCGTCTTCTTCATCCACATCACTGTAGTCGCTGTTCTCCGAATGAGAGT	62213
3238	Qy	CGGTTTCGTCTGCGGTTTTGGAGTAGTGGAGGGGAATCTTGGTCAAAAAATCCTCAG	3297
62214	Db	CGGTTTCGTCTGCGGTTTTGGAGTAGTGGAGGGGAATCTTGGTCAAAAAATCCTCAG	62273
3298	Qy	GTGCAGTCATGTCAATCACATCGGGGGCGCTCCGTACCCGGCGGTACCGCAACCCCGTAC	3357
62274	Db	GTGCAGTCATGTCAATCACATCGGGGGCGCTCCGTACCCGGCGGTACCGCAACCCCGTAC	62333
3358	Qy	GGGACCCGGCGAAACAAAAACGTCCTGTTCTCTTTTTTCTAGTGCCTCCGGAAATCGGCA	3417
62334	Db	GGGACCCGGCGAAACAAAAACGTCCTGTTCTCTTTTTTCTAGTGCCTCCGGAAATCGGCA	62393
3418	Qy	GCATCTTGGAGTGGCGCGGGGCTCTTGACCGGTCGGATAGAACATAGCATAGCCGCG	3477
62394	Db	GCATCTTGGAGTGGCGCGGGGCTCTTGACCGGTCGGATAGAACATAGCATAGCCGCG	62453
3478	Qy	AAACGTTTCACCTGT-AAAAACGCACTGCCTATCCCGATCAGTTAATATTTCAGTCAACGG	3536





QY 6775 AAAAAATTTTAAACAAACTCGCCAGGTAAGTAAACGATGGTACGACAGGTCTGTGAAA 6834  
 Db 65726 AAAAAATTTTAAACAAACTCGCCAGGTAAGTAAACGATGGTACGACAGGTCTGTGAAA 65785  
 QY 6835 AGTGTCTATCATCTCTTACCTCCCTTGAAGGTTTGGTATACACGACCATGTGAGA 6894  
 Db 65786 AGTGTCTATCATCTCTTACCTCCCTTGAAGGTTTGGTATACACGACCATGTGAGA 65845  
 QY 6895 TAGGCACCAAGCTCAGGGGGAGGTGCGATCGTGAAGGGGTAAACAATTCGTGGCGGTGGGG 6954  
 Db 65846 TAGGCACCAAGCTCAGGGGGAGGTGCGATCGTGAAGGGGTAAACAATTCGTGGCGGTGGGG 65905  
 QY 6955 GTACTCCAGTTTAAAGTCTCCGCGAGGACGCGTACGAGTTCTTGTGAAATCTTACGCA 7014  
 Db 65906 GTACTCCAGTTTAAAGTCTCCGCGAGGACGCGTACGAGTTCTTGTGAAATCTTACGCA 65965  
 QY 7015 GCCCGCATCCACAGGGGAGTACTGAGTTAAAGTCCCGAAATCTTATCGTTGTTAAT 7074  
 Db 65966 GCCCGCATCCACAGGGGAGTACTGAGTTAAAGTCCCGAAATCTTATCGTTGTTAAT 66025  
 QY 7075 TTTGTGAGCTCCTTTAGACAGTAAACCATGTTGCTCCGACACTGCGGTACGATGAACCG 7134  
 Db 66026 TTTGTGAGCTCCTTTAGACAGTAAACCATGTTGCTCCGACACTGCGGTACGATGAACCG 66085  
 QY 7135 CTCTCTTGTCTGTGTCATGAGCACCTGGAGTATCTTAGACAGGTAAGGCGCGGTTTC 7194  
 Db 66086 CTCTCTTGTCTGTGTCATGAGCACCTGGAGTATCTTAGACAGGTAAGGCGCGGTTTC 66145  
 QY 7195 CACTCTAAGGGTTCCCGCTTGTGTAGACGGGATTTGGCGGATTTGACGGTCAGTGTGTCAT 7254  
 Db 66146 CACTCTAAGGGTTCCCGCTTGTGTAGACGGGATTTGGCGGATTTGACGGTCAGTGTGTCAT 66205  
 QY 7255 TTCCAGGACTCATCTCCGAGTGGCGGATTTGGAACTTTGATCATATGATTTAAAGATG 7314  
 Db 66206 TTCCAGGACTCATCTCCGAGTGGCGGATTTGGAACTTTGATCATATGATTTAAAGATG 66265  
 QY 7315 GTCTTTCCCGTAATGGCGACTCATGTGCTGTTTAAACGTTTCAAGGGTGTAAACAACT 7374  
 Db 66266 GTCTTTCCCGTAATGGCGACTCATGTGCTGTTTAAACGTTTCAAGGGTGTAAACAACT 66325  
 QY 7375 CAGAGGTCTGTGTGAGCAACTTTTGGCGACTCTCCGAAACGCTCAGGGATAAGCTGTA 7434  
 Db 66326 CAGAGGTCTGTGTGAGCAACTTTTGGCGACTCTCCGAAACGCTCAGGGATAAGCTGTA 66385  
 QY 7435 CCAGCGGTGCTTCTTATAACGAAGCATACGTCAAAAAACAATGACCCACGTCACCTAAA 7494  
 Db 66386 CCAGCGGTGCTTCTTATAACGAAGCATACGTCAAAAAACAATGACCCACGTCACCTAAA 66445  
 QY 7495 TAAGCTCACAGAAAGTCTGTGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCAT 7554  
 Db 66446 TAAGCTCACAGAAAGTCTGTGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCAT 66505  
 QY 7555 CTCGGGCTCATCGCGGATCTGTAATCTTTAAACCTGTTCTGCTCTACCGGGGTTCCCG 7614  
 Db 66506 CTCGGGCTCATCGCGGATCTGTAATCTTTAAACCTGTTCTGCTCTACCGGGGTTCCCG 66565  
 QY 7615 GGTATAAACCCCGGGCCCGCACTGTAAACGTCCTCGGTGGCGAGAAATGGGCCCAGGGCAT 7674  
 Db 66566 GGTATAAACCCCGGGCCCGCACTGTAAACGTCCTCGGTGGCGAGAAATGGGCCCAGGGCAT 66625  
 QY 7675 CGTGAGAAATCTCACAGAGGGGCCCTTTGTTGTAACCGAAAGATGTTTATACGCTCTGC 7734  
 Db 66626 CGTGAGAAATCTCACAGAGGGGCCCTTTGTTGTAACCGAAAGATGTTTATACGCTCTGC 66685  
 QY 7735 CTGACGAGGGGTCTTTATACCGCCCGGAGTTGGCGAGAGTTCTGCAAGAGTCTACCGGGA 7794  
 Db 66686 CTGACGAGGGGTCTTTATACCGCCCGGAGTTGGCGAGAGTTCTGCAAGAGTCTACCGGGA 66745  
 QY 7795 AATGAAGGCCAAGTGTGTGGGGGCTGGCGTCTCATATGTTGACGGGGCTTATCAT 7854  
 Db 66746 AATGAAGGCCAAGTGTGTGGGGGCTGGCGTCTCATATGTTGACGGGGCTTATCAT 66805  
 QY 7855 GGCTATTGCCGACTCGGTCTCTGCTACTTATAATACCTCGATGCCAGGGGAAAACTGGA 7914

Db 6806 GGCTATTGCCGACTCGGTCTCTGCTCACTTATAATACCTCGATGCCAGGGGAAAACTGGA 68665  
 QY 7915 ACTTAGACTCAAAAGCACTGTGCAAACTTTGTTTTTCAACCCACTTTTCTTCAAGAAATCTTT 7974  
 Db 68666 CCTTAGACTCAAAAGCACTGTGCAAACTTTGTTTTTCAACCCACTTTTCTTCAAGAAATCTTT 68925  
 QY 7975 AGCCCTCTATGAGACTATGCGCCAAACGGGAAAGATGGTTCCTGACAAATTTATTTTACCATCAC 8034  
 Db 68926 AGCCCTCTATGAGACTATGCGCCAAACGGGAAAGATGGTTCCTGACAAATTTATTTTACCATCAC 68985  
 QY 8035 CGGTACGGCCGAGAAAGAGGCGCCTCTGCTGACTGGAAGTACTAGCGGGGATGACGTGTCC 8094  
 Db 68986 CGGTACGGCCGAGAAAGAGGCGCCTCTGCTGACTGGAAGTACTAGCGGGGATGACGTGTCC 67045  
 QY 8095 GGGAAAGCAGCCTTGTCCCGGATTCCTTAATCTCTGCCAGTATGCGAAACGGGGGCTTCTCCC 8154  
 Db 67046 GGGAAAGCAGCCTTGTCCCGGATTCCTTAATCTCTGCCAGTATGCGAAACGGGGGCTTCTCCC 67105  
 QY 8155 GGCAACCCCTGTTGACCTCAGTAATGCTTAGAAAAATCCAGAAATCATCTCTCAGCGGCCCC 8214  
 Db 67106 GGCAACCCCTGTTGACCTCAGTAATGCTTAGAAAAATCCAGAAATCATCTCTCAGCGCCCC 67165  
 QY 8215 ACCCTGAGTCAATTTGTCATCAAAACACGCAACCCAGTCTGCTCAGTCAGTCAGGAT 8274  
 Db 67166 ACCCTGAGTCAATTTGTCATCAAAACACGCAACCCAGTCTGCTCAGTCAGTCAGGAT 67225  
 QY 8275 TATTACGCCAACCCAGGGCGTGTTCGCCCAATGTTTATGACACAGTGGAAAGCGGT 8334  
 Db 67226 TATTACGCCAACCCAGGGCGTGTTCGCCCAATGTTTATGACACAGTGGAAAGCGGT 67285  
 QY 8335 GTACAGAGCAATTCACCAACAGGACAGAGCGCTTATTTTGGCCCGCGCACATACCGGTTTC 8394  
 Db 67286 GTACAGAGCAATTCACCAACAGGACAGAGCGCTTATTTTGGCCCGCGCACATACCGGTTTC 67345  
 QY 8395 GACATCTGCGGCCCTTGGCCCGCATATCGCATGTTTCCCGAGTTCGCCGACAGTCTCCGCA 8454  
 Db 67346 GACATCTGCGGCCCTTGGCCCGCATATCGCATGTTTCCCGAGTTCGCCGACAGTCTCCGCA 67405  
 QY 8455 GGTGGAAGGGTCCCGGGCGCTCGATTGCGGGAACCCAGCATGCGTTCCGACGCCGCGTT 8514  
 Db 67406 GGTGGAAGGGTCCCGGGCGCTCGATTGCGGGAACCCAGCATGCGTTCCGACGCCGCGTT 67465  
 QY 8515 ACCGCCCAATGTCCCGCGCAAGAGATGGAACCGGTAGCACAGTGGGAACCGTCCCGT 8574  
 Db 67466 ACCGCCCAATGTCCCGCGCAAGAGATGGAACCGGTAGCACAGTGGGAACCGTCCCGT 67525  
 QY 8575 AAAAAACGTCACATCGGAGCGCGGTATACGCTCCACTGTTTAAATATACCAATAATAGA 8634  
 Db 67526 AAAAAACGTCACATCGGAGCGCGGTATACGCTCCACTGTTTAAATATACCAATAATAGA 67585  
 QY 8635 CTTAACTGCTCCCGTCAGGGTCCGGCCAGAGTCCGGCCGATATCGCAACACTCCAGAGTC 8694  
 Db 67586 CTTAACTGCTCCCGTCAGGGTCCGGCCAGAGTCCGGCCGATATCGCAACACTCCAGAGTC 67645  
 QY 8695 CCGCATGCGCGCGCTCTCCGCCCTTCCGCCGAAACCGCCGAAACCGGTCCCGCTAAGAG 8754  
 Db 67646 CCGCATGCGCGCGCTCTCCGCCCTTCCGCCGAAACCGCCGAAACCGGTCCCGCTAAGAG 67705  
 QY 8755 AAAGCAGCCACCGAGGACGTTGGGAGACAAAGACTGAAAGGGGACGTTTGGGGGGCGCGC 8814  
 Db 67706 AAAGCAGCCACCGAGGACGTTGGGAGACAAAGACTGAAAGGGGACGTTTGGGGGGCGCGC 67765  
 QY 8815 AACAGTAAACCACTTTCCCGGACCGTCCGGGATGCGGTTTCCGAGCAGAGGGCTTAT 8874  
 Db 67766 AACAGTAAACCACTTTCCCGGACCGTCCGGGATGCGGTTTCCGAGCAGAGGGCTTAT 67825  
 QY 8875 CGATTATAATCGAAAGCTCCAGGATGTAAACCGCGAAACGCAATCTGGAACCAAAACGACGA 8934  
 Db 67826 CGATTATAATCGAAAGCTCCAGGATGTAAACCGCGAAACGCAATCTGGAACCAAAACGACGA 67885  
 QY 8935 CGACATGCTAGCGGTATCTTACAGGACTGTATGAGTACAGTCCCGCCCGGCTATCGA 8994

Db 67886 CGACATGCTAGCGGCTATCCTACAGGACCTGTATGGACTACAGTCCCCCGGCCCATCGA 67945

Qy 8995 TTCCCTCCCTCCAGCAACTCGGCAATGAGGAGATATTTCCAGAGGTAGTTCGCGCCATCTAG 9054

Db 67946 TTCCCTCCCTCCAGCAACTCGGCAATGAGGAGATATTTCCAGAGGTAGTTCGCGCCATCTAG 68005

Qy 9055 CGGCCACGATCGCCTTGA 9073

Db 68006 CGGCCACGATCGCCTTGA 68024

RESULT 3  
AF2411163  
LOCUS  
DEFINITION  
Cercopithecine herpesvirus 17 immediate-early transcription factor  
Rta mRNA, complete cds.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AF2411163.1 GI:13506778  
Cercopithecine herpesvirus 17  
Cercopithecine herpesvirus 17  
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
Gammaherpesvirinae; Rhadinovirus.  
REFERENCE  
AUTHORS  
TITLE  
Lin.S.F., Robinson,D.R., Oh,J., Jung,J.U., Luciw,P.A. and Kung,H.J.  
Identification of the b2IP and Rta homologues in the genome of  
rhesus monkey rhadinovirus  
Virology 298 (2), 181-188 (2002)  
22127395  
12127781  
2 (bases 1 to 1945)  
Lin.S.F., Robinson,D.R., Oh,J., Kung,H.J. and Luciw,P.A.  
Cloning and characterization of cDNAs encoding Rhesus Rhadinovirus  
(RRV) ORF50  
Unpublished  
3 (bases 1 to 1945)  
Lin.S.F., Robinson,D.R., Oh,J., Kung,H.J. and Luciw,P.A.  
Direct Submission  
Submitted (02-MAR-2000) Biological Chemistry, School of Medicine,  
UC Davis, 4645 Second Avenue, Rm. 2400, Sacramento, CA 95831, USA  
FEATURES  
source  
Location/Qualifiers  
1. 1945  
/organism="Cercopithecine herpesvirus 17"  
/mol\_type="mRNA"  
/isolates="17577"  
/specific\_host="Macaca mulatta"  
/db\_xref="taxon:154334"  
65..1798  
/note="ORF50"  
/codon\_start=1  
/product="immediate-early transcription factor Rta"  
/protein\_id="AAK28330.1"  
/db\_xref="GI:13506779"  
/translation="MDTDDNGVKKLRSCVDNFCGLSPRLDKLYHALLINEAYVK  
KDDPRDLNLTVEVCHRCIVMECASLPIISGLIADLNLENLCLYRGSVKTRGAATN  
VPCAECAGIVRLITERALCTTEKFIASCSGVVTPQLARVLHDVNAEMKAKCLGA  
WRLLICCRPIMAIADSVLTNDABGKIDRLKALCKVQFPIFQRIILAPMLL  
ANGKMPDNYFTTGTAEKRPVVTGSGMTCPGSSLPDLSLILFVCEPGLLPAPLV  
DLSNVLNEPILISAPLSQFVITNTHPSLPQSIVITPTQGVFQCQFMDTWKAVSQ  
SIHOACTPILAAALTGSTSAAPGHIACSPVAGTSRQVSGAGVDCGKPCVQOPAL  
PNPVAKMETVAQLGNAFVKNVHIGRVYVAPLVNPIIDLTSPSGSGOSPADIANTP  
ESRMAGSPPPAEATATVAKSKQREDVADKELKGVREGAATVNHPRPGSGMEVRE  
QGFLDIESTDVITANASGPKNDMDLAILQDLVLQSPFPAIDSPSSNSDNEEIFPE  
VSPFSGHGSP"

ORIGIN  
Query Match 18.9%; Score 1714.4; DB 14; Length 1945;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7358 CAGGGGTATAAAATCTAGAGGTCTCGGTGGACAACTTTTGGGACTCTCCGAAACCG 7417

Db 83 CAGGGGTATAAAATCTAGAGGTCTCGGTGGACAACTTTTGGGACTCTCCGAAACCG 142

7418 CTGAGGGATAAGCTGTACCAAGCGGGTCTTCTTATATAACGAAGCATACGTCAAAAACAT 7477

Db 143 CTGAGGGATAAGCTGTACCAAGCGGGTCTTCTTATATAACGAAGCATACGTCAAAAACAT 202

Qy 7478 GACCCACGTGACCTAAATAAGCTCACAGAAAGAGTCTGTGATGTATTGTAAATGGAATGT 7537

Db 203 GACCCACGTGACCTAAATAAGCTCACAGAAAGAGTCTGTGATGTATTGTAAATGGAATGT 262

Qy 7538 GCCTCCCTGGGTCCCATCTCTCGGGCTCATCGGGATCTGAATCTCTTTAACTGTTCTGC 7597

Db 263 GCCTCCCTGGGTCCCATCTCTCGGGCTCATCGGGATCTGAATCTCTTTAACTGTTCTGC 322

Qy 7598 CTCTACCGGGTTCCTCCGGGTAAAAACCCGGGGGCCGCCACCTGTAAACGTCCTCCGTGGCA 7657

Db 323 CTCTACCGGGTTCCTCCGGGTAAAAACCCGGGGGCCGCCACCTGTAAACGTCCTCCGTGGCA 382

Qy 7658 GAATGCGCCACGAGCATCTGTGAGAAATCTCACGAGAGGGCCCTTTGTTGTACCGAAAG 7717

Db 383 GAATGCGCCACGAGCATCTGTGAGAAATCTCACGAGAGGGCCCTTTGTTGTACCGAAAG 442

Qy 7718 ATGTTTCATAGCTCTGCTGCGGGGTCTGTATACCGCCCGGCTGGCCAGAGTTCTG 7777

Db 443 ATGTTTCATAGCTCTGCTGCGGGGTCTGTATACCGCCCGGCTGGCCAGAGTTCTG 502

Qy 7778 CACGACGTCTACGCGGAAATGAAGCCCAAGTGTGTTGGGGCGTGGCGTCTCATATGT 7837

Db 503 CACGACGTCTACGCGGAAATGAAGCCCAAGTGTGTTGGGGCGTGGCGTCTCATATGT 562

Qy 7838 TGCAGGGGCTATCATGGCTATTGCGACTCGGTCTCTGCTCACTTATATACCTCGAT 7897

Db 563 TGCAGGGGCTATCATGGCTATTGCGACTCGGTCTCTGCTCACTTATATACCTCGAT 622

Qy 7898 GCCAGGAGAAATCGGAACCTTAGACTCAAAAGCACTGTGCAAACTGTTTTTCAACCCATC 7957

Db 623 GCCAGGAGAAATCGGAACCTTAGACTCAAAAGCACTGTGCAAACTGTTTTTCAACCCATC 682

Qy 7958 TTCTCTTAAGAAATCTTAGCCCTATCGACTACTGCGCAACGGAGAGTGTCTCTGAC 8017

Db 683 TTCTCTTAAGAAATCTTAGCCCTATGACCTACTGCGCAACGGAGAGTGTGTTCTGAC 742

Qy 8018 AATTATTTCATCATCGGTACGGCGAGAGAGGGCGCTCTGCTGACTGGAAGTACT 8077

Db 743 AATTATTTCATCATCGGTACGGCGAGAGAGGGCGCTCTGCTGACTGGAAGTACT 802

Qy 8078 AGCGGATGACGTGTCCGGGAAAGCAGCCTTGTCCCGATTCCTTAATCTGCGCAGTATGC 8137

Db 803 AGCGGATGACGTGTCCGGGAAAGCAGCCTTGTCCCGATTCCTTAATCTGCGCAGTATGC 862

Qy 8138 GAACGGGGCTTCTCCCGGCACCCCTGTTGACCTCAGTATGCTTTAGAAAATCCAGAA 8197

Db 863 GAACGGGGCTTCTCCCGGCACCCCTGTTGACCTCAGTATGCTTTAGAAAATCCAGAA 922

Qy 8198 ATCATCTCTCAGCGCCCAACCCCTGAGTCAATTTGTCTCAAAACACGACCCCGTGTG 8257

Db 923 ATCATCTCTCAGCGCCCAACCCCTGAGTCAATTTGTCTCAAAACACGACCCCGTGTG 982

Qy 8258 CCTCAGTCAGTCAGATATTACGCCAAACCCAGGGCGTGTGTTCCCGGCCAATGTTTATG 8317

Db 983 CCTCAGTCAGTCAGATATTACGCCAAACCCAGGGCGTGTGTTCCCGGCCAATGTTTATG 1042

Qy 8318 GACACGTGGAAGCGGTGTACAGAGCATTCACCAACGAGGACAGAGCCCTATTGTTGCC 8377

Db 1043 GACACGTGGAAGCGGTGTACAGAGCATTCACCAACGAGGACAGAGCCCTATTGTTGCC 1102

Qy 8378 GCCGACATAACCGGTTTCGACATCTGGCGCCCTCTGGCGCGCATATCGCATGTGTTCCAGTT 8437

Db 1103 GCCGACATAACCGGTTTCGACATCTGGCGCCCTCTGGCGCGCATATCGCATGTGTTCCAGTT 1162

Qy 8438 GCCGCGAGCTCTCGCGAGGTGGAAGGTTCGCGGGCGTTCGATTGCGGGAACACGACATGC 8497

Db 1163 GCCGCGAGCTCTCGCGAGGTGGAAGGTTCGCGGGCGTTCGATTGCGGGAACACGACATGC 1222

JOURNAL MEDLINE PUBMED REFERENCE	(human herpesvirus 8) genes in Kaposi sarcoma Proc. Natl. Acad. Sci. U.S.A. 93 (13), 6641-6646 (1996) 96270595 8692871 4 (bases 122794 to 123567; 123809 to 127297; 129372 to 130400; 130551 to 134441)
AUTHORS	Ceserman, E., Nador, R.G., Bai, F., Bohenzky, R.A., Russo, J.J., Moore, P.S., Chang, Y., and Knowles, D.M.
TITLE	Kaposi's sarcoma-associated herpesvirus contains G protein-coupled receptor and cyclin D homologs which are expressed in Kaposi's sarcoma and malignant lymphoma
JOURNAL MEDLINE PUBMED REFERENCE	J. Virol. 70 (11), 8218-8223 (1996) 97048116 882957 5 (bases 1 to 137508)
AUTHORS	Moore, P.S., Boshoff, C., Weiss, R.A. and Chang, Y.
TITLE	Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV
JOURNAL MEDLINE PUBMED REFERENCE	Science 274 (5293), 1739-1744 (1996) 97094384 8939871 6 (bases 1 to 137508)
AUTHORS	Russo, J.J., Bohenzky, R.A., Chien, M.-C., Chen, J., Yan, M., Maddalena, D., Parry, J.P., Peruzzi, D., Edelman, I.S., Chang, Y. and Moore, P.S.
TITLE	Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8)
JOURNAL MEDLINE PUBMED REFERENCE	Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14862-14867 (1996) 97121480 8962146 7 (bases 1 to 137508)
AUTHORS	Russo, J.J., Bohenzky, R.A., Chien, M.-C., Chen, J., Yan, M., Maddalena, D., Parry, J.P., Peruzzi, D., Edelman, I.S., Chang, Y. and Moore, P.S.
TITLE	Direct Submission
JOURNAL	Submitted (17-OCT-1996) Dept of Pathology/Genome Center/Division of Epidemiology, Columbia University, 650 West 168th Street, New York, NY 10032, USA
REFERENCE	8 (bases 1 to 137508)
AUTHORS	Russo, J.J., Bohenzky, R.A., Chien, M.-C., Chen, J., Yan, M., Maddalena, D., Parry, J.P., Peruzzi, D., Edelman, I.S., Chang, Y. and Moore, P.S.
TITLE	Direct Submission
JOURNAL	Submitted (02-MAY-1997) Dept of Pathology/Genome Center/Division of Epidemiology, Columbia University, 650 West 168th Street, New York, NY 10032, USA
REMARK COMMENT FEATURES	Sequence update by submitter On May 3, 1997 this sequence version replaced gi:1718251. Location/Qualifiers
source	1. .137508 /organism="Human herpesvirus 8" /mol_type="genomic DNA" /db_xref="taxon:37296" /cell_line="BC-1; virally infected primary effusion lymphoma derived cell line" /note="sequence derived from overlapping lambda phage and cosmid clones isolated from libraries; A prototype sequence for the terminal repeat (TR) unit is deposited under GenBank Accession Number U75639. The left junction variant TR sequence unit immediately preceding base pair 1 of the LUR is deposited under GenBank Accession Number U75700. The remaining LUR sequence extending from this sequence to the right TR unit is not available due to difficulties in cloning and sequencing of this region"
misc_feature	1. .137508 /note="long unique region LUR"
CDS	105. .974 /note="ORF K1" /codon_start=1 /protein_id="AAC57081.1" /db_xref="GI:1718252" /translation="MFVYVCSLAVCFRGLSLSLSSPNLCQGVSTPTVLTCLSNAL SLFSWICNTRLLRTERVLDTACNFTCVQSGHRQS1WTRPAQVLTCLCAQ

PSNVTGCGHVTLYCSTGNNVTVMLPGRNETVSYQKYNYFTLMSOTEGCYTCSNG  
LASRLNRICFCARCANITPEHTVSVSTTGFRLTSLNLSKIKIHAITROVVVKEA  
KSTHFIEVHFLVMTVALIGTMCIGILGIIIFALCQKORDSNKTVPOQLQDIYSLHD  
LCTEDYTPQVDMT"  
1142..2794  
/note="ORF 4: The HVS ORF 4 homolog has alternatively  
spliced messages encoding membrane bound and secreted  
proteins; complement binding protein vCBP homolog"  
/codon\_start=1  
/protein\_id="AAC57082.1"  
/db\_xref="GI:1718253"  
/translation="MAPLRQTWLMTFTWVTGQNEKKSQKTLGYRLKMRGDGIIA  
VSTELRCRSGYTYARNITATLCGGTSEPTATCNKKSCEFNQVQNGKVIPIHGG  
QDALICANISVYNEGVFLVGYRYVYCMIGASQMAWSSPPFCEKEKCHRPKIN  
GDFPKDQYVYNDVAFECNEGYTLVGHISACAVNTWTSMPTCELAGCKFFSVT  
HGVPIQGFSLTYKHQSVTFACNDGFLVSGSPITCNTEVDPPLPKVLELIDDPNN  
SNPGRLHPNPEKPNVQFQSNYVETPPKPEDTHTAATCDNCEQPKKILPTSEGN  
ETTSNIIKQLEDEKTIQOPNTHITSALTSMAKGNFNKNTNSDHLIASTPTSDQ  
DAPSPVQVPTNNTNAPTRILTSIHIEGFSNITSSEKATSSLSHSHNSHNDGGI  
YTNLTKQTPSTNKTNSQAKSTKPRVETHTNKTNSPAISLTSADVPQRPREPTL  
PPIFRPPASKNRYLEKQLVIGULLTAVALTCGLITLFIHYLFFR"  
3210..6611  
/note="ORF 6: ss DNA binding protein ssDBP homolog; EBV  
BALF2 homolog"  
/codon\_start=1  
/protein\_id="AAC57083.1"  
/db\_xref="GI:1718254"  
/translation="MALKGPQTLLENIGSAAPTGPGYLYAYLTNFPITGEASLLNG  
YPEAKVSLPLHGLTVESDPLNKAHVKKIDATVASVKLTSHYREAIVFNTHLFO  
PIQGGLEKLCRESBELPGEFTFVQHQHGLTWSEACQLPCANIEPMAVIVTEGF  
KERLYGKLVVPSQTPPHIGEHQAFKPLVDEDLFGPSRAQELCRFNPDISLHLY  
DSIFTGAALRVSTVIQASERQFVHDQYKIPKLVQKQAFQPCASRGTDGSLMY  
IDSLAELNGSYGLSPTEQSCVLYNDTWPIFENCETPDARLAREVHAEQALH  
IGALFAANSVLYTVAKLPQKQNGDANMYSFYLOHGLGLYSEATYKENGASAFK  
GVFSDALDSSGYTLOHVASPSPHLLARMCVYLOPLPHKNTNSQSYNVVYVGT  
APQSMCDLQCGQCPACINTFYRMKDRFPVLSNVKRPDYVITVGTAGYNDLEILGN  
PATERRESEGNVEDAPKYTYWOLCONTEKLASMGISEGDAALRLTVIDLPSFVKY  
FKGLDSTVEALLKFNCKMIKNYNFRENISKVHHLQPCACNVQACPVEFLIYK  
SLUTVTDICLTCEMTYQNPVAGIVPSENLQHFQMTWPKFACGPFGLITYGEL  
KIVHGFQSDLRDITDAIGMFAPEARMQVRIARMLVPKTIKIKRNIIFSNTGAE  
IQAGFMKPAQDSYTVGGPYKFNALHKLTPSTKTSALYLMHKIGQTTKRNPLPG  
VSCHEITELCNVYKASQAFEEINVLDDVPLDTLSYAKILNSSLIRACGQTFYAT  
LSCGLSPVLOVAPAEVPHVPLVGLSSPDEYRAKVAGRSVTIVQSTLKOAVTNGRLR  
PITVTVLVNKTNGNTNVHFCANLGVSGRGVDNRLRPESVPPKKNVSSMLKR  
HVTWPLVRLVRIIVGINSGEFEAEAVRSQVNLDRDNENLAKTVVLELVLPWR  
SSCASITBEDVYILGFYAVLGVDEVLSTLSTVGQGVPTWAEVAVSIQDIIIDDCLELQ  
FVGFPEPCLTQGSVVVEELFPSPGVPSLTVGKRRKIASLLSLDL"  
6628..8715  
/note="ORF 7: transport protein homolog; EBV BALF3  
homolog"  
/codon\_start=1  
/protein\_id="AAC57084.1"  
/db\_xref="GI:1718255"  
/translation="MAKELAAYVADVSAALAMDCLLSYADPATLDTKSIALITGKFQ  
LHGTLLPLLRQNAHBCSLGLEHFWKTLMLWPRNECALAENCLQKSIFFSCIVT  
QHATSNRSVRFNWNWALELKLINDVEIIFPKLSVYFICIGSGALIGLELWLF  
VGLRGTSPVPGDLYVSNLPLECIQVCLTENQCTSLQAMLPDPACSHICTPACGE  
PVGLRLENELKQLQLOTPRSIPTPCQSVRQDDLEIRQSLMAVGDHITFGVTRSVL  
EISNLIYSSGSDATCDGDRDCSHLASLTHEADMEKRVLDLAGCLGERGPKHPFD  
CFRPSLETFLCGSLFSSVEDTIESLQKDCSSAFYQOYNTTALQKVDYVRLSKDL  
AAGLNKSTESQSEARQVGGKPEEVLRAKHQELYLQKVARDDGPKKSDC  
IRHGHILSOTLGLRWGVSIVNEASALQNHFLRAQFISLPQDLTVDCTPFRFENSK  
YKNSLYCQRLGHEVEILTFEYKLITGSLSKRHTLFPSPNVTLQCFEAGMLPH  
QKMWSEMIPIEPEDNIEPNQVSEFNQDINHLOKRAWEYIRELVLSLYNRT  
WERELKILLTPQSGPGEFEPKPAGLTGLYLTETSAPLVLDVKKYGMIPKDIYALLY  
HHQLSNHNSQV"  
8699..11236  
/note="ORF 8: glycoprotein B gb homolog; EBV BALF4  
homolog"  
/codon\_start=1  
/protein\_id="AAC57085.1"  
/db\_xref="GI:1718256"  
/translation="WTPFRSLATIGTVILLVCFCAHAHSRGDTFTQSSSPFPQSSS

CDS

CDS

CDS

CDS

KAPTGPKEEASGPKSVDFQYFRVCSASITGELFRFNLEQTCPTDKYKHQBGLLVYK  
KNVPHIPKVRRYRKIATSVYVRLGLESATINKYELPRFVPLVYIEIASHMDSTYQCFSS  
MKUNVANGVENTFDRDDVNTTVELOPVEGLTDNIQRYFSOPVYVAEPCWPGIYRVRT  
TYNCEIYDMTARSABPNYFYVSLGDTVEVSPCYNESSGSTTSPKNGSLVQVVLNH  
TVVTSRGSPTPKRIFIVETGATYILSWASEKTYTAVCBIAKTKPSTQTHQDHS  
FHFVANBITATITAPLTPVANFDITYSCLSDINILWASKAKULASHVNGTVQYEH  
TTGGLYLWQPMASAINLTHAGDSNGTSPSPASQVMDTSSASRKRKESASTAAAGG  
GSTNLSTYQLOFAYDKLRDINGVLEELSRQVMDNLMWYELSKNPTSVMTA  
IYGRPVSAKFGDAISVTECINVDQSSVNIHKSRLTNSKDVCPARPLVTKFLNSNL  
FTGOLGARNBIILLNNOVETCKTCEHYFITRNETLVKYDYAVLRTINTTIDISLNTF  
IAMLNFIQNDFAKIELYSSAEGKLASSVPDLTFMREYNYTHRLAGLRDLDNDTI  
DNKMERVRDLSEIVADLGGKTVNVASSVTVLCSLATVGFNFKHPGLGMLMII  
IVTALLIIFMLSRNTWIIAQPVMYIYIPVDRAAPSPSGAGAPFREEIKNILLGNHQLQ  
QBERQKADDDUKKSTSPFORTANGRLRGKLPLOTSLDISPEIG"  
11363..14401  
/note="ORF 9: DNA polymerase homolog; EBV BALF5 homolog"  
/codon\_start=1  
/protein\_id="AAC57086.1"  
/db\_xref="GI:1718257"  
/translation="WDFNFPIDTRGCPNTRVROPTSPSPVTPSETRVCRLLIPACF  
QTGRPGVAVDITTFPTTYGPKRGEVFAGETSIIWKRGRQARNAPMHSLLIHFVVD  
IVETTYADRCEDVFFSFQTDIIPSGTVLKLGLRTLDGASVCNVFQRCYFYFLAQ  
GVNLTHVLOQALQAGFGRASCSTFEPVRKKILRAYDQYAVQKILTSSEPMRTLS  
DRLTTCGCEVFESNVDAIRRVLDHGFSTFGWYBCSNPAPRTQARDSWTELEFDCSWE  
17.7%; Score 1607.6; DB 14; Length 137508;  
Query Match  
Best Local Similarity 51.6%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 3869; Indels 342; Gaps 21;  
QY 1 ATGAGAGATTCCGTCGGATGACCAACACACGTCCGAAACCAATCCGGGGTTTCATCTTGAAC 60  
DB 64992 ATGACAGCTCGGAAGGTGCACGTATATGACAGAGCCAGCCCGGTTTCATCTCAAC 64951  
QY 61 ATGAGCTCCGATGCCAAAGTCAGGGGTGTCGTGATGATCATCGTCAGTCGCTGTCAAAATA 120  
DB 64952 ATGACATCGGATGCCAAAGTCAGGAGCGTAGTGAACACAGATGACAGGTTTCAAAATC 65011  
QY 121 ACTACACGCCACCGGAATCGGTTGTACACCTGGCTTCCGATCCGGCTGAAGACATCC 180  
DB 65012 ACTACCTCGCCCGAGATGGCTGGTATGACCTAGTTCGATCCACTCGAAGACGAA 65071  
QY 181 GGGCGGTTCTTTGCCGTTTACCGTTTATCTAATAACGGGAATCTGTGTGTGGGAAAGT 240  
DB 65072 GGGCCCTTTCTGCCGTTTTCGGCATACGTATAACGGGAGATGTCAGGAGCGGGGAAAGC 65131  
QY 241 ACCAGCATATCGGCCCTGTACAAATTAACCTGCTGATCACGGCGGCGCACCAACATA 300  
DB 65132 ACCAGCGTATCGGCCCTACATCAGAACTCACTCACTGCCCTAATACGGGGCTACAGTGA 65191  
QY 301 GCGGCACAGAACTATCGCGTCGCTAAAGACGTTTCTGTCCACGATCTTCAGCGCTTTT 360  
DB 65192 GGGGCACAGATCTTTCCAGGGCTTTTAAAGCTTACTGTCCACTATATACACGCTTC 65251  
QY 361 GCGTTTAAAGAGCGGACACATCAATATAGCCGTGAGAAAGCTCATCAGACCGGACCGTA 420  
DB 65252 GGAATTCAGAGACGACACATTAATATCTGCGAGAGGAAAGTCCCAAGGTAACTCAGTCC 65311  
QY 421 TCCATAGAGCAAAATTCAGCAACAGAGGCTATCGAAGTATTGGCGGGTTTATAGTGGACAT 480  
DB 65312 TCCATCGAGCAACTCCAGAGATACAGCTGGCTAGTACTGCGCAACTGTCCCGATATT 65371  
QY 481 ATGAAGAAGGTTATGGCGAAACCAACCCAAATGGCATGTACGGGATCTATATCCACCGCAAT 540  
DB 65372 ATTCGAGAATTTATGCGAAGAAACAAAGGGGCGAGTATAGTCCCTCTCTCAAGCGCT 65431  
QY 541 TTTGAAACCCCTCGAGATACCGGACCGGTGTTTATGAGCTTCCAAATATTATTGTAATC 600  
DB 65432 TTCAGACTCTTTTGGCGTATGGTGGAGCCAAATTTGTGACAGGTAAACATATTCTGTATA 65491  
QY 601 GACGAGGCGGAAACCCCTGTCTCTTACATACCTACCAACCGTGTGTTCTTTTACTGGTTC 660  
DB 65492 GACGAGCTGGAACCCCTCTCTGCTCCATATTATTGACGGCGGTGTTCTTCTTCTATTGTT 65551

CDS

Query Match

Best Local Similarity

Matches 4481; Conservative

QY 1

DB 64992

QY 61

DB 64952

QY 121

DB 65012

QY 181

DB 65072

QY 241

DB 65132

QY 301

DB 65192

QY 361

DB 65252

QY 421

DB 65312

QY 481

DB 65372

QY 541

DB 65432

QY 601

DB 65492

QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTTACCGCCAGGGGGCGGTTCCTGTCATAGTATGC 720  
 Db 65552 TACAACAGTTGGCTGACACCCCGCTATACAGAAATGGTGGCTTGCATAGTCTGC 65611  
 QY 721 GTCCGTTCCGCCAACGAGCAAAACGCGTTCAGTCAACTTAACAACGCGGAGCAAAAG 780  
 Db 65612 GTGGGCTCTCCACCCAGACGCGAGCGCTTTTCACTGGGTCTTCAACCAACAGCAGCAGAGA 65671  
 QY 781 ACGGAGATATCGTCGCGGAGAACATCCAAATTCATGATAGGGAAGAAGTCTATCT 840  
 Db 65672 AACGAGATATCTCCCTGTGATATATGTGCTCACTTCTTATGGGAAACGTGAGGTTGA 65731  
 QY 841 GAGTACGTACATTTGGAGAGAACTGGGCGCTGTATTATATAAACAAGCGCTGACCGAT 900  
 Db 65732 GATTATATTAGGCTGACAGAAATGGGCGCTATTATATAAACAATAAGCGCTGTACGAT 65791  
 QY 901 CTCAGTTCCGACACCTCTTAAATTTTAGAGTAAATCTTCCCACTTCTCCGACGATC 960  
 Db 65792 CCCAGTTTGGTCACTTGTCTGAAGACCTTAGAATATTAATCTAGACHATATCACAGATTA 65851  
 QY 961 ATGAGTTACGTACAGAGATTTGTCTTCTTAAAGTAAGATATATGATCTTTTAGAATAC 1020  
 Db 65852 ATGACTATATAGATAGTTTGTGTTCCGAAGATAGATTTCTGGACCGCTCGAGTAT 65911  
 QY 1021 ATTGGCTGGACCGATCTTTTGTTCATAGCGAGGTAAAGCGTATCTACAAACCTA 1080  
 Db 65912 GCAGGCTGGACAAAGACTTTTCACTCACACGAGGAGTGAAGTCTTTTCTGGCAAGCTG 65971  
 QY 1081 CACACATGTCTAACGCTAGGGGCGGATACAGGGACACGAAGCTCTTTTACCTGTCCCGTG 1140  
 Db 65972 CACACCTGCTGCGAGTAATAGGATGTGTGTCCACAAAGCTTTTCACTGCGCCAGTG 66031  
 QY 1141 GTGTGCGAGGTGTTTGTGAAGCGGTTTGAGGAATAAAGCGGCGCGTCAACCTCACCAAC 1200  
 Db 66032 GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGCGGCGTAGGCTCTCACAC 66091  
 QY 1201 CTCACGTCACGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTTTCGAGTTT 1260  
 Db 66092 ATGACTCCCATGAATGGGTAAACAAATATCTTTTTCAGGCTAAGTAACTACTCGAGTTT 66151  
 QY 1261 GTGGACAGGACATGTCATAGTCGCCACGGAATCCACTGAAGCTTTCGACCCAGGTTACC 1320  
 Db 66152 GCTGATCAGGACATGGCTGTGGTTGGACCTATATCACAGACGCGTCCACAGATCAC 66211  
 QY 1321 TTTATCACCAAGTTGTTTAAACACGCGCTGCTTTTAAATGGAAACCAAAAATGC 1380  
 Db 66212 TTGCCACTAAATTTGTCAAAAACAGCTATGCTACCCCTTACTGGAAGACCAAAAATGT 66271  
 QY 1381 ATATGTGGGTTTCAGGTAAGTACTTTTGAAGTTCAAAAGAAATCCTAGACAGCGAATCTTC 1440  
 Db 66272 ATATGCGGTTTTCAGCGGTCAATACCAAGATTCAAGTCCATCTAGACGGGAGCTATT 66331  
 QY 1441 GTGGAGACGATTCGCAAGACCGTCCGGAATAGTATACGTTTCTTAAACAGCTACTG 1500  
 Db 66332 ATCGAAAGTCAATTCGCAAGATAACCCCGCTTATGTGTACAGTTTCTTATGACCTGCTA 66391  
 QY 1501 TACAACGCCATGTAATCGTTTTCAGCGTACGGGTGACTAGGTCAACAGAGAAATATCTG 1560  
 Db 66392 TATATGCCATGTAATCTATTTAGCGCAAGGGTGAAGCAGGGCATGAAGATTTCTC 66451  
 QY 1561 CAAGACTAAATTTTGCAACCTCCCGCGCTCTGCGAAACCGGGGTGTAGACCTTCAA 1620  
 Db 66452 AGGACCTCAGGAACTCGCGGTGTCTCAAGAGGTGATCTCTGAGATGAGCTCCGAGGAC 66511  
 QY 1621 ACGGTTGTAAGAGTTTAAACCTGGAAGACGACATCTTACACAGCTGTGCACTGCCCG 1680  
 Db 66512 GTTCTGGGCGAGGGGGAACAAGATG-----CCTTCTACTCCCGCAGCTCCCA 66565  
 QY 1681 CCACCCCGGGGTATCACTCTCCCTCCAGTTTGTGACACGTAAGTCTGCGCCCTTAAAGGAC 1740  
 Db 66566 CCATCCCCCAACCCAGCGGCTCTTCCAAACACTGTGTGCTTATTAATCTCCGGGCGCAAGGA 66625  
 QY 1741 GTGTTGCGCTCCAGAAATAAGGTGGCGTGTGCTGTGTTGCGGGAGTTTGAGAAGGA 1800

Db 66626 CTAATCTGCAACAGCTGGCCCTGGCACGCGACACTTTTGGTACAGAGTTCTCTCCACTCC 66685  
 QY 1801 ACGTTTTCGCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTT 1860  
 Db 66886 GATTTTTCACAGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTT 66745  
 QY 1861 TCAGAACGTCTCAACCGGCTGTTTGGCGTTTGCATCGACCGTTTGAATCGTATAAATAAAG 1920  
 Db 66746 TCCCGCGGCTCCACGCTTAGTGGCATAGCCATCCACTATAGACACTATATAATCCAG 66805  
 QY 1921 GGGTACACGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTACCGTTTAC 1980  
 Db 66806 GGAATATAGTTTCTCCAGTGAAGATTCGCGCGTCCAGGAGACAGCCCTTACGCGAGAC 66865  
 QY 1981 CTCAGGAAGATGCGCTCCCTCGTCTGAGGACTCTAGCGGTTTATCGCGTGCCTA 2040  
 Db 66866 CTGCGCAAGATGCGCTCCCTCCATAGTTTGTCCAGACTCATCGGGTTTCAITTCGCTGCTG 66925  
 QY 2041 GAGATTAACATAACCAATTTGACCGAAACCATGGAGACGCGGAGCTTTTCCAAAGTGTGC 2100  
 Db 66926 GAAATAACGTCACCAAGATGACAGACCCCTCGAAGTGGCGAGCTGTTTAAACATATGT 66985  
 QY 2101 TGTCCGGGACTATGGGTCAGCTCAAAATTTAGCCATGACCATCGTAAAGGCAACAGGA 2160  
 Db 66986 TGTCCAGGGACTACGCTATCAGTTCTAACTGCTATGACCATAGTGAAGGCAACAGGG 67045  
 QY 2161 ATGTCGTTGGAGCGAGTAGCGGTAGTATTTGGTCCCAAGAACGCTCCAGACAGCCAC 2220  
 Db 67046 GTTTCACTAAGTGGGTGGCCATATCGTTCGGCAACCAACCGCAATATCAGAGCCAGTCTA 67105  
 QY 2221 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAAATTTATTTGGTTCATGGACAGCAACCCC 2280  
 Db 67106 GTGTATGTGGTGTATTCAGGGCCCATCGACGCTGTTACCTGTTAATGGACAGTAAATCCC 67165  
 QY 2281 CTTAAACCCCTCTCAGAGAACAGTGTGATTAACACTCCCGCAAGCATATAGTCCGCGCC 2340  
 Db 67166 CTTAAGCTAATGGACCGCGGTGACGCCAGTCCCGCATCTCTCAAAGTACATCATCAAGGCC 67225  
 QY 2341 CTCACACACCAACCAACCCCTCATCTACTAA-----ATAAATAAACAAGCAAAATGGAAAAATTTGTG 2373  
 Db 67226 CTATGCAACCCCAAGACTACTCTGATCTACTGACCCCGTACCCCTCTCTTAGGACACTGAT 67285  
 QY 2374 -----ATAAATAAACAAGCAAAATGGAAAAATTTGTG 2406  
 Db 67286 GTGTTTGGGAATAAAGCATGAGACTTGACACCTATAATGGTCTGTATTTGACACCATTTCT 67345  
 QY 2407 TTTTATTTCACTCAACCAACCGCCACGGATAGTTGTCATTTCCACACACACCGGGGGGTGG 2466  
 Db 67346 TTTATTTCACTCCAGCCACCGCATATATGACCGCTTTCCACACAGGGGTGGCGTGG 67405  
 QY 2467 CGGCCATAGTTTGACACCAAGAACCGCGGCGCTCTGGGTTAGAAAGTAGACCC----- 2521  
 Db 67406 AGGCCAGGATCGGGTTGGGTGGTGGCTGACCTGGACCCCGCGGTAGTTGTGCTTCTGATG 67465  
 QY 2522 -----GCCCGCGGTGGTGGGTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 2564  
 Db 67466 AAATCGAGTGGCGGAAGTACTGGGAGATTGGGTTGGAGGTGACCTTTTGTGCTGACG 67525  
 QY 2565 GTACCAACATAAATCAAGTCTGATCAACGTTGAGTCCCTCGTGCAGCGCGCTCTTTG 2624  
 Db 67526 GAGACAGCATCAAGCTACGCGGACGAGGAGTCTCTGCTGCTGTCTACTCCCGAGGATA 67585  
 QY 2625 CAGGCGGACCAACCGCGCTGTGACGTTGATGCTG-----CAT 2664  
 Db 67586 TAAATATCACGAGACCCACTGCTTTTGGCGCTTAAAGTTGGTGTCTGTGGAGCGCAACA 67645  
 QY 2665 CGTGTGCTGTCTGAGATGAACACAGGTCAATCAGTGTGCGCGCGCTGCGCTCAT 2724  
 Db 67646 CATCTCTGCTACAGAGAGCGGTAGACTGCTTTTGGCTTCTGCGCCACGCTCATGA 67705  
 QY 2725 CGCGCTCTGATTCGPTCCCATGTCTCAACAGAGCTCGGTGCTGCTGCTGCTGCTGCTGCTG 2784



Db 67706 GCCGATCTCTGACTCAATCTCCGCTTGGTCTTCTCCGCTCTCCGACGAGGGTG 67765  
Qy 2785 AAATAATAATGGTCTCTTGTCTGCTGAGTTGATCCAGCAACACAGGTTCTTCTTCATCGG 2844  
Db 67766 GCTGGTGGAAAAATGGCGCGTCCGTAAGACCGCGCTCATTTGTTACGTCGGAGAGT 67825  
Qy 2845 AAGACCATCCGCGCTCTCGGCTGATCTAGATATATACCTGACACAGCGCGGCGAGGCC 2904  
Db 67826 TGGAACTGTCTATCGGTATCAGAGTCCGATGTCCAGATCCGCTGCGGTGGCGCGC 67885  
Qy 2905 G----- 2905  
Db 67886 GCAGGGGGGCCACGAGGSCCTTCATCAGGGTGCCTGTATGGTGAACCTTTGTGTTCCAG 67945  
Qy 2906 -----TGCGAAAGCGGTGAGTCTTCCGCCACTTTTCATCAGTGT 2946  
Db 67946 GTACACTATTTCTGAAGCAGGTGAAGTCGATATGCCCGGTCCAGTGTATGCCGCCA 68005  
Qy 2947 CCTCGTAGTCATATAAAAAATCGTCTTCTGCTCTGTGATTCAGATGACTGGGGTGGCG 3006  
Db 68006 TCGGTTCCAGGATAGACNACCCCTCGTCTGTGAAGGTGAGAGCCACGAGGGGAATC 68065  
Qy 3007 GGTGTTGGCAGTAACCCAGCAAGCGCGCG-----GACACACAAACCG 3033  
Db 68066 CGTCATCTGACTAACCCATCCATGGAGCCTCGGACTCCGCGGTGCTCGTTGAAGTGC 68125  
Qy 3054 GTCACTGGCGCGCGCTCGGTGGTGGGACCTGTGAAAGTAATGACGCTAGGT 3113  
Db 68126 GCACGGGCGCGCTACCACTGTACCGGTTGGGGGTATGGCGCGTCTGCCAGAGGCC 68185  
Qy 3114 AGCGCGCGCTCTGGTTAAAG-----ACTCACGCTCCCGTG 3150  
Db 68186 TCGGCGCAAGTGAGATAAGGTTGAAAAAGTCTGCGGGTACCCCTCGCTCGTCTT 68245  
Qy 3151 CGTGGGATCTATTCTAGTGGGAGGAGACCGCGGTTCTGCTCTTCATCCACATCAC 3210  
Db 68246 CCTCTGGAACATGCTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 68305  
Qy 3211 TGTAGTCTGCTTCTCCGAATAGAGTCCGTTTCTGCTGCGTCTTTTGGAGTAGGGAG 3270  
Db 68306 ATTGCGCGCTCGACTGATCCGGGATATCTGTAGATCCAGAGGGTGTGTCGGCGCGATG 68365  
Qy 3271 GGAATCTTGGTCAAAAACATCTCAGGTGAGTCAATGATCACTGCGGGGCTCCGT 3330  
Db 68366 CGGTGCTCTCGGAGAGCTGCTGCGGCGAGATATCTATCACCGTGGGTCCAGCAT 68425  
Qy 3331 ACCCGGGGCTACCGCAACCCGCTACGGGA-----CCCGCGCAACAAAAAC 3378  
Db 68426 AGCGCGCGCTGCGCAATCTGAGTGATGAAGAGGTGGAGGTGGGATATGAACT 68485  
Qy 3379 GTCTGTTCTCTTTTCTAGTGGCGCGGAAATCGGAGCATCTGGAGTGGCGCGC 3438  
Db 68486 TCACGGGGGTCTCTGCGAGCGCTCTTCAATTGGAAGCATTTCTCTTCATCGTGTG 68545  
Qy 3439 GGTCTCTTACGGGTCCGATAAGAC-----ATAGCCATGGCC 3476  
Db 68546 TGTAGACAGGTCTCAAAACATGCCATGGCCTGTACGGGTTGACCGTGGAGGC 68605  
Qy 3477 GAAACGTTCCACTGTAAAAACGCACTGCTATCCCGATCAGTTAATTTCCAGTCAACGG 3536  
Db 68606 GGAATTTCAAGCACAGGTTATTGCTTTTACTGCTCCCAACAGGCCCGAGTCCACAG 68665  
Qy 3537 CACCCCTCGATTCTGAACTAGATGATCTTGCAGTTAAATGTTGAGCCCTAGAA 3596  
Db 68666 TCTACCGCGGTGGCGAGTCAAAATAGTCTGTGGTAGGTTAAAGTATACAGCCCTGGA 68725  
Qy 3597 ACTGCGGCGAGGAGAACCGGTTGCGCGTGGCGGTAGCCAGGGGAGCAGATGGCGCG 3656  
Db 68726 ACCGAGGCCATCGGAGTGTGGCCACCAAGAG-----GCCAGCGAGATGATGATCTGGG 68782  
Qy 3657 CCTGTAGGACGAGATGTCTGTCTGCTTATCAGCGTCCGCTTGTCTATAGCCTTTCTTC 3716  
Db 68783 CCGTAAGCACCAAGTGTCTGTGCGCTTTATGAGCGGAGTTCTGTCAATGGGCTTGGCC 68842

Qy 3717 CCACAGCATAAAAACGACCTCTGAAGCTTGGCAGAAAGGACGCTTATATCTAGCTCG 3776  
Db 68843 CCACAGGAGAAAAACGCAATGTTTAACTTTGAGGATATGCTACTGATGATGAAACTCG 68902  
Qy 3777 TGAACCAATCCAGCAAGGTTGGCGTGTGACCCCGCTTCCCTCTTTTCCACCGTCAGAA 3836  
Db 68903 TGAACCAATCCAGCAAGTCCCTCGTGTGAGCGCGCCCTCCCTCTCTCCACCGTCAAAA 68962  
Qy 3837 TGTGTTTAAAGCAGAACTCCCGGTTTGGCCCAACAGTCTAAGCACCCGTGAGAGGAG 3896  
Db 68963 CTGTGTTTAGTACAAACACACCTTGGCAGCCCGAGCTGTGAGGCAACCCGTGGAAAGGAG 69022  
Qy 3897 CGGTGAACCCAGGTACAGTATTGGCTATCTCTTAAATAATTTTGAAGCTTTGGAGGTA 3956  
Db 69023 TACTGAAATTTGGGAGCGGAAGCCTCTAGCTCTCTAAAGATGCTTCTCAAACTGGTGAA 69082  
Qy 3957 TACTGTAGTCCGAGCAACGCTAAAGCTAGTCCGGTGGCTTGAACCGCGATGTTAGGAT 4016  
Db 69083 CCTGACATTCGGATCCACACTAAACGCCAGGCGAGTAGCTTGGCCCTTGTGTAGGGGT 69142  
Qy 4017 CTTGGCCAGGATCAGACCTTTAATATCCCTCGGTTTCGCAACAATAGACACACACATAA 4076  
Db 69143 CTTGGCCCTAAGATCACCACTTTAATATCTCTGATTCGAGGAGTGGGACCAACCATCA 69202  
Qy 4077 TTCTATCTATCGGTGATNAATCACGGTTACGTTGCTCATATCCATAACCGGTTGAGCA 4136  
Db 69203 GCTTGTCTGTGGGGATACACTGTGTGGTGTAGCTTAAGTTCCGAAATCTGTCTGAGCA 69262  
Qy 4137 GAGCGCGACTCTCTGTTTAAATGGAGAAAGATTAAATTTCAAGGCCATGAGTCGC 4196  
Db 69263 GCAGAGCAGTTTCTGTTTCAAGAAATGATGAGAGGCTCAGAAAGGAAATCCACTTAGGTG 69322  
Qy 4197 TAAAGTAGCA-----ACGTTTGTGTGGAGGTTCTTTCTAAACT 4234  
Db 69323 CCAGTAACAGATCCCGTCTGCCACCCCTGACTGATGATAGGTGCTCCCTAAAGACCG 69382  
Qy 4235 TCTGTGACATTTTACTCCGAACTATTTGTTTAAAGCAACCCCTCCATAAATCCCGCGTC 4294  
Db 69383 TCTGTTGCAACCATGGGTCATGTTGAACTATTTTCCCTTTTGAACCTGCGTGGCGCTCTC 69442  
Qy 4295 TGGTAATTTACTTTACGTATACGTGAAACTTTTCTGCTGTGAGAAATCCAAACAGAG 4354  
Db 69443 CGGCTGCTGTTTTAGCCCGAGTCTGACTTCCGCTAACGAACCTGTCCGTTTATGCGCC 69502  
Qy 4355 GTGCGCGGTATAGAGTTCCCGTTTCACTACCGTTTATGTGAAAGT-----CT 4404  
Db 69503 TTTCCACAGCTTATTATAATTATGTTTACGTTGTGAATAGAGCTATCTGCAGTGGTCGCG 69562  
Qy 4405 TCAAACTCGGCTAAABACTCATCAAGACTATATTGCATATAATTCAACAATTTGAAGCTCT 4464  
Db 69563 TTAACACTAGTATAGGCGGTCAAACTTCTGTTTGAATCCACACACACCTCAGGTTT 69622  
Qy 4465 TCTCTGGGACACTGTAAATTTAACTTGTACAAATCCTTATAAAAAACGCCCTTAAGTGA 4524  
Db 69623 TCCGTGACGCCCGAGGACCCCAATCTTCGAAACGACCGCGACTAAAAATGACCTCAGATTA 69682  
Qy 4525 AATCCATTTACATATTTCTGTGGTAGAGGATCAGCTTTATATCGCAATGTTGCTGTA 4584  
Db 69683 AACCATTTCAGCATGTTTCCACGGTAATTCGCGCTGTTTGTCTTCCAGCTTGGCTATA 69742  
Qy 4585 CACAGCCCTGGAACATAAGCTAAATTTGTTTAAATGAAAAGAACCAATCTCAACCGGTCC 4644  
Db 69743 CAGACCCCGTTGCACTGATTTCCGATCGCGAGTGGATAGAGTGGACCGCAAGAACAAAC 69802  
Qy 4645 GGAATAATAAGTCACTAAGTGTATTTTATCAACATGATTTTACTACAAAGCCCGTGA 4704  
Db 69803 GGCAGGGTAGAGGTGCGGATGCTGAAATTTGCGCAACATGTTAAGGCGAGTATGCGGTGA 69862  
Qy 4705 CAACATGTAATTAATAAAAAATCCACATGTAATAATAAACTTAAGGTATACAT---ACTT 4761  
Db 69863 GATGTGACCAATAGGGTGTCCACAGGACCGCAATAGCGCAAGATCCCCCATGGGCAA 69922



Db 72081 ACAAATGGAGGAGGACCACTCTGTGACAGTCCGGGCTCTGGGTTCTGATTTCCAGGGGC 72140  
 Qy 6944 TGGCGCTGGGGGTACTCCAGTTTTAGGATCTCCGGCAGGACCGCTACGAGTTCTTCGTCA 7003  
 Db 72141 GCCGTGTGGGGTATTGGAGAGTCAAAACTCTGGGCAGTCCCTTAATGAGCTCTCTCA 72200  
 Qy 7004 AATCTTACGAGCCGCATCCACAGGGCAGTACTGAGTTAATAGTCCCGGAATCTA 7063  
 Db 72201 AAACCTATGAGCCAGCGTCCACTAGTGGCAGATGCGGTTAATAACACCCCTTATCTTG 72260  
 Qy 7064 TCGTTCGTAATTTTGGAGCTCTTTAGACAGTAAACCATGTGTCGACACTCGCGT 7123  
 Db 72261 TCGTTCGCAAGTTGTACAACTCTGCAGGAAATAGCCAAATTCGCCCTAGCCGGGA 72320  
 Qy 7124 AGATGAACCGCTCTCTCTGTGTGTTGTCATGAGCACTGAGATATCTTAGACAGGTAT 7183  
 Db 72321 ACCAGTACGGGTCTGCTTTGTGCGGTGTCGGAACAATATCTGAATGCTCTTTTCAAGGTAT 72380  
 Qy 7184 AGGGCCCGTTTCACTCTAAGGTTGCGCTGTGTAGCGGATTCGGGCATTTGACGTC 7243  
 Db 72381 AGGCTCTTCAAGTTTGTAGCGGGTACGTGCGAGTCTGGAATGAGGGTGGCGAGAC 72440  
 Qy 7244 AGTGTGTCATTTACCGCACTATCATGCCAGTGGCGGATGTTGAACTTTGACATATGA 7303  
 Db 72441 AGGGTATCTAATCTCTGAAATATCTGATCCAGGACGGGTAATGATACCTAAACAGATGG 72500  
 Qy 7304 TTTAAAGATGCTTTTCCCGTAATGGGCACTCATGGTCTGCTTTTA-----AC 7352  
 Db 72501 TTGAACAGGTGATCTTTAAGGGGCTTCTCGATGTCATTTGTAACATATGACGCCAC 72560  
 Qy 7353 GTTCACAGGCTGTAAAAAAGTCTGAGGCTCTGAGTGTGACAACTTTTGGGACTCTCGG 7412  
 Db 72561 TCTCTCTTGTAGGTAAAGCTTCGCGGCTCTGTGTGGAAGCTTCGTGCGGCTCTCGG 72620  
 Qy 7413 AACCGTCAAGGATAGCTGTACACCGGCTGCTTTTATTAACGAGATGACTCAAAA 7472  
 Db 72621 AGGAATGAAGGCCCACTCTACAGTGTGCTCTTATAATGACGATACGAAACAA 72680  
 Qy 7473 AACATCACCACTGACCTAAATAGCTCACAGAGAGTCTGTCGATGTTTAAAGTGG 7532  
 Db 72681 TCTACGATCCAGTACCTAAATAGAGTGTGGAAGTGTGCTCGATTTGAAAG 72740  
 Qy 7533 AATGTCCTCCCTGGGTCCTATCCGGGCTCATCGGAGTCTGAATCTTTTAACTGT 7592  
 Db 72741 AATGTTCCAAAGTGTGGTGGCTATGTCGTCTTTACAGACATTAACATGTTTAACTTT 72800  
 Qy 7593 TCTGCTCTACCGGCTTCCCGGTTAAACCCGGGGCCGCCACCTGTACGTCCTGCT 7652  
 Db 72801 TCTGCTTTTCTGCTCTCTGAAATGAGGACCAAGGCGCGGCTACACGTCGCTAT 72860  
 Qy 7653 GCGCAGATCGGCCCAAGGCTATGAGAAATCTCACAGAGGGGCTTTGTTGTACCG 7712  
 Db 72861 GCGCAGGAGCATCCCAAGGCAATTTCTCGGATCTCTACGGAGAGGATCTTATCTGCACAG 72920  
 Qy 7713 AAAGATGTTTACGCTGTGCTGAGCGGGGTGTTATACGGCCCGAGTTGGCCAGAG 7772  
 Db 72921 AAAGGCAATTTCTGACAGCGCATGACGCGGGGTGAGCTGCTCCAGCCATATGTAAGC 72980  
 Qy 7773 TTCTGACAGGCTCTACGCGGAATGAAGCCAAAGTGTGTTGGGGCGTGGCGCTCTCTCA 7832  
 Db 72981 TACTACAGGAATATACCTGAATGAAGCCAAATGCTCTGGGGCTGGAGGCGACTCG 73040  
 Qy 7833 TATGTTGAGGCGGCTATCATGGCTATTGGCACTGCGTCTGCTGCTATATATACCC 7892  
 Db 73041 TCTGCAATCGGAGGCGCCATATGATATTAACCTCTTCCCTACTGAAGCTCTCAACACGT 73100  
 Qy 7893 TGGATCCGAGGGAAGTGAACCTTAGACTCAAGCACTGTGCAAACTGTGTTTTCAC 7952  
 Db 73101 AGATACCGCGGGTCTCTGAGCAGTCCAGGCGCTCTGCTTTGTTTTCAC 73160  
 Qy 7953 CCATCTTTCTAAAGATCTTAGCCCTATGAGCCTATGCGCACTGCGCAACGGGAAGTGGTTC 8012  
 Db 73161 CGGTCTACCTCCGAGGATTAAGCGCGCTGAGATCATGACCAAGGGTCACTGCGCC 73220

RESULT 5

U93872  
 LOCUS U93872 133661 bp DNA linear VRL 09-JUL-2001  
 DEFINITION Kaposi's sarcoma-associated herpesvirus glycoprotein M, DNA replication protein, glycoprotein, DNA replication protein, FLICE inhibitory protein and v-cyclin genes, complete cds, and tegument protein gene, partial cds.  
 ACCESSION U93872  
 VERSION U93872.2  
 KEYWORDS GI:14627174  
 SOURCE Human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus - Human herpesvirus 8)  
 ORGANISM Human herpesvirus 8  
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.  
 1 (bases 29032 to 30108; 117733 to 118431)  
 Zhong, W., Wang, H., Herndler, B., and Ganem, D.  
 Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genes in Kaposi sarcoma  
 Proc. Natl. Acad. Sci. U.S.A. 93 (13), 6641-6646 (1996)  
 96270595  
 8692871  
 2 (bases 17242 to 17856)  
 Neipel, F., Albrecht, J.C., Ensser, A., Huang, Y.Q., Li, J.J., Friedman-Kien, A.E. and Fleckenstein, B.  
 Human herpesvirus 8 encodes a homolog of interleukin-6  
 J. Virol. 71 (1), 839-842 (1997)  
 97138401  
 8985427  
 3 (bases 123309 to 124082)  
 Li, M., Lee, H., Yoon, D.W., Albrecht, J.C., Fleckenstein, B., Neipel, F. and Jung, J.U.  
 Kaposi's sarcoma-associated herpesvirus encodes a functional cyclin  
 J. Virol. 71 (3), 1984-1991 (1997)  
 97184528  
 9032330  
 4 (bases 122660 to 123226)  
 Thome, M., Schneider, P., Hofmann, K., Fickenscher, H., Meinl, E., Neipel, F., Mattmann, C., Burns, K., Bodmer, J.L., Schrotter, M., Scaffidi, C., Kramer, P.H., Peter, M.E. and Tschopp, J.  
 Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors  
 Nature 386 (6624), 517-521 (1997)  
 97242415  
 9087414  
 5 (bases 1 to 133661)  
 Neipel, F., Albrecht, J.C. and Fleckenstein, B.

Cell-homologous genes in the Kaposi's sarcoma-associated  
rhadinovirus human herpesvirus 8: determinants of its  
pathogenicity?  
J. Virol. 71 (6), 4187-4192 (1997)  
97296220  
PUBMED  
REFERENCE  
AUTHORS  
Neipel, F., Albrecht, J.-C., Ensser, A., Huang, Y.-Q., Li, J.J.,  
Friedman-Kien, A.E. and Fleckenstein, B.  
The genome of human herpesvirus 8 cloned from Kaposi's sarcoma  
unpublished  
6 (bases 1 to 133661)  
Neipel, F., Albrecht, J.-C., Ensser, A., Huang, Y.-Q., Li, J.J.,  
Friedman-Kien, A.E. and Fleckenstein, B.  
Direct Submission  
Submitted (17-MAR-1997) Virology, University of Erlangen,  
Schlossgarten 4, Institut für Klinische und Molekulare, Erlangen  
91054, Germany  
8 (bases 1 to 133661)  
Neipel, F., Albrecht, J.-C., Ensser, A., Huang, Y.-Q., Li, J.J.,  
Friedman-Kien, A.E. and Fleckenstein, B.  
Direct Submission  
Submitted (09-JUL-2001) Virology, University of Erlangen,  
Schlossgarten 4, Institut für Klinische und Molekulare, Erlangen  
91054, Germany  
Sequence update by submitter  
On Jul 9, 2001 this sequence version replaced gi:2246466.  
Location/Qualifiers  
1. 133661  
/organism="Human herpesvirus 8"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:37296"  
/notes="sequence derived from overlapping lambda phage  
clones isolated from two libraries of SauIa partially  
digested DNA from Kaposi's sarcoma biopsies"  
105..959  
/notes="ORF K1"  
/codon\_start=1  
/protein\_id="AAB62619.1"  
/db\_xref="GI:2246494"  
/translations="MFYVVCISLAVCFRGLISLOSSNLCPGVISTVTLTSPNT  
SLPTSYCNDTRLRLITQPTFTDLICNFCVCGQGHSHLWITVYPOVPLQTLGQ  
LSRTYCGQVITLYCSGNVTVWHLPGNRTVSTQYKYNFTLNLQWEGCYACNSG  
LSRLNRCVFCARCANLTPEHTVSSTSTPRTSTNRGTTDLQVLWKEAKSTNL  
HQVHFVFLVALLGTMCGLIGLTIFAHCQQRDSNKTVPQQLQDYSLHDLCLCTED  
YTPQVDVWY"  
1127..2779  
/notes="ORF 04, complement control protein homolog"  
/codon\_start=1  
/protein\_id="AAB62602.1"  
/db\_xref="GI:2246477"  
/translations="MAFLRQLTWLMTFTNVIQDNKEKCSQKTLIGYRLKMSRDGIIA  
VGTSLRCSRGTSTYARNRTATCLOGTWSPTATCNKSCNPGEIQNGKVIIFRGG  
QDALGANIYVCNCGYFVLGVEYVYCMIGASGVAMSSSPFCEKEKCHRPKFN  
GDKPDKAYVENDAVHFCNEGVTLVGPHSIACAVNNTWTSNMPCELAGCKFPST  
HGVIQGFSLTYKHOSVTFACNDGVFLRGSPITICNTWBDPLPKCVLEIDDDNN  
SNFGRLLPRTNEXGNVPSVNTPTPPTPEHTHTAATCDTNCCEQPKLPTSEGN  
ETIISNITIKQLEDKTSGPNTHTISALTSMAKAGNFNTKNSDTHLHASTPSQD  
DATPSVPTPNNTNAPTSLTSLHIEGPNSTISEKATSLTSHNSKNDTGGI  
YTLNKTQTPSTNKSQAKSTKPEVTHNTKNTSPAISLTSADVPQRPBRTPL  
PFIPRPAPSKRYLERQLVGILLTAVALTCLGLTTLPHYLFPR"  
3194..6592  
/notes="ORF 06, major ssDNA binding protein homolog"  
/codon\_start=1  
/protein\_id="AAB62603.1"  
/db\_xref="GI:2246478"  
/translations="MALGKQQLTEENIGSAAPTGPCGVLYALTYHREPIGEASLILNG  
YPAKVPSPFLGLTVESDFPLNVKAVHKKIDATVASKVLYTHREAIYFHNTHLFG  
PIFGKGLKLEKLSFGFTVEQGHGKIDATWSAPQLPCANEIFWAVITVIGF  
KSLTYGKLVFVPSQPTTVHIGEHQAFILYDDELFGPRAQELCRFPNPDLSRILV  
DSIYTGIAQLKVDVSTYQASERQFHKDQYIKPLIQVARDPQCASTGDSGLTM  
IDSLVABLGNYSGLSFTFEGQDSCVLNVTWPIFENCEFTDARLEVYHVAEQALH  
IGAOIPLANSVIVITRAKTIPOKNROGDANNYSFYHLGHLGYSIATVKEGASAFK  
IGAOIPLANSVIVITRAKTIPOKNROGDANNYSFYHLGHLGYSIATVKEGASAFK

CDS		LPHLVYKQARBELQIHDRIPVFDVADPGLSRSLASHPEYVKQHLRVAVDLY PKLHVAVANIQCFLFQNNISATVAMLNFLDIPVTPFP"	
		/note="ORF 10, herpesvirus saimiri ORF 10 homolog, conserved in other gamma-herpesviruses"	
		/codon_start=1	
		/protein_id="AAB62594.1"	
		/db_xref="GI:2246469"	
		/translation="MQTEATFILGDWEIVTSNCRCTCSSLTGCGPLRSGDYTRLRIP	
Query Match		17.7%; Score 1602.8; DB 14; Length 133661;	
Best Local Similarity		51.5%; Pred. No. 0;	
Matches 4478; Conservative		0; Mismatches 3872; Indels 342; Gaps 21;	
Qy	1	ATCGAGAGTCCGTCCGATGGACCAACACACGTCGACCAAAATCCGGGGTTCATCTTGAAC	60
Db	65191	ATGGACAGCTCGGAAGGTGCACTGATATGGAGAGCCGAGCCCGGGTTCATCTCAAC	65250
Qy	61	ATGACGTCCGATGCCAAAGTCAGGGGTGTCGTGGATCAGCTCAGTCGCTGTCAAATATA	120
Db	65251	ATGACATCGATGCCAAAGTCAGGAGCGTAGTGAACAGATTGACAGAGTTGTCAAATATC	65310
Qy	121	ACTACAGCCACCGGAATGGTGTGAGACCTGGCTTCGATCCGGCTGAAGACTCC	180
Db	65311	ACTACATCGCCCGGAGATGGCTGGTATGACCTAGAGTTGATGCCATCGGAAGACGAA	65370
Qy	181	GGCCCGTCTTCGCGTTTACCGTTTATCTAAATTACGGGAACCTGCTGGTCTGGGAAAAGT	240
Db	65371	GGCCCGTCTTCGCGTTTTCGGCATACGTAATTAACGGGACTGCGAGACGGGGAAAGC	65430
Qy	241	ACCAGATATCGGCCCTGTACCAAAATTTAACTGCTGATCAGCGGGCCGACACCATTA	300
Db	65431	ACCAGGATATCCGCCCTTACATCAGAATCTCAACTGCTTAATTACGGGGGCTACAGTGTA	65490
Qy	301	GGCGCAGAACTATCGCTCGCTGCTAAAGAGCTCTGTCCCAAGATCTTCAGCGCTTT	360
Db	65491	GGCGCAGAACTTTTCAGGGCTTTAAAGTCTTACTGTCCCACTATATACCAAGCTTC	65550
Qy	361	GGCTTTAAGAGCCGACACATCAATATAGCCGTGAGAAAAGCTCATCAGACCGAGCCGTA	420
Db	65551	GGATTCAGAGCAGACACATTAATATCTGCCAGAGAAAGTGCCCAAGTAACTCAGTCC	65610
Qy	421	TCCATAGAGCAATTCAGACACAGAGCTATCAAGTATGGCGGTATAGTGGACATT	480
Db	65611	TCCATCAGCAATCTCCAGAGATACGAGCTGGCTAGGTATCGGCCAACTGTACCCGATAT	65670
Qy	481	ATGAAGAGTTTATGGGAAAAAACCAATGSCATGTACGGGACTATATCCAAAGCGAAT	540
Db	65671	ATTCGGAATTTATGCCAGAAACAAAGGGCAGTATAGCTCCCTCTCAAGCGCT	65730
Qy	541	TTTGAAACCTCTCGAGAAATGACGGAGCGTGTATGACATTCGAAATATTATGTAATC	600
Db	65731	TTCAAGACTCTTTCGCGTATCGGTGGAGCCAAATTTGTGGACGAGTAACATTATCGTGATA	65790
Qy	601	GACGAGCCGGAACCTGTCTCTTACATACTCACACCGTCTGTGTCTTTTACTGCTTC	660
Db	65791	GACGAAGCTGGAACCTCTGTCCCATATTTTACGGCGGTGTGTCTTCTATTGTTT	65850
Qy	661	CTAAACAGCTGGCTPAATAACCCCTTTTACGGCAGGGGGCGTTCCGTGCATATATGC	720
Db	65851	TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTCCGCTTGTGCATAGTCTGC	65910
Qy	721	GTCCGTTCCGCAACGAGACAAACGGCTTCCAGTCAACTTACAAACACGCGGACGCAAAAG	780
Db	65911	GTGGGGTCTCCCAACGAGAGCGCCCTTTCAGTCGGTCTTCAACACAGCGAGCAGAGA	65970
Qy	781	ACGAGATATCGTCTGCGAAGACATCCTAACTTATGATAGGGAGAGGTCGTATCT	840
Db	65971	AACGAGATATCTGCTGTGATAATGTGCTCACCTTCTTATTTGGGAAAAACGTCAGGTTGCA	66030
Qy	841	GAGTACGTACATTTGGAGAGAACTGGCGCTGTTTATAAACAACAGCGCTGCACCGAT	900
Db	66031	GAATTAATATAGGCTGGACGAAATTTGGGCCCTATTATATAAACAATAAGCGCTGTACGGAT	66090

Qy	901	CTCAGTTCCGACACACCTTCTAABAAATTTTAGAGTATATATCTTCCATTCTCTGACGAAGTC	960
Db	66091	CCCCAGTTTGGTCACTTGGCTGAGACCTTAGAATATATCTAGACATATCACCAGAGTTA	66150
Qy	961	ATGAGTTACGTAGACAGATTGTCGTTCCATAAGTAAGATTATGATTCCTTTAGAATAC	1020
Db	66151	ATGGACTATATAGATAGGTTTGTGTTCCGAAGAGTAAGAATCTTGGAACCCGCTCGAGTAT	66210
Qy	1021	ATTGGCTGGACCCGACCTTTTGTACATAGGAGGTAAAGGCTATCTTAAACAACCTA	1080
Db	66211	GCAGGGTGGACAGACTCTTCACTCACACAGAGGTGAAGTCTTTTCTGGACACGCTG	66270
Qy	1081	CACACATGTCTTAACGCTAGGGGGCGATACCAGGGACACGAAAGCTTTTACTGTCGCGTG	1140
Db	66271	CACACCTGCCTGTCGAGTAATAGGATGCTGTGTCCACAAGCTTTTCACTGCCACGTG	66330
Qy	1141	GTGTGCGAGGTGTTTGAAGCCGTTTGAGGAATACAGGGGGCGTCAACCTCACCAAC	1200
Db	66331	GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGGGGCGTAGGCTCACACAC	66390
Qy	1201	CTCACCGTGACCGAGTGGGTGACAAAAAAGCTCTTTAAGTTAAGTAATTTATTCGAGTTT	1260
Db	66391	ATGACTCCCATAGATGGGTAAACAABAATCTTTTCAGGCTAAGTAATCTACTCGCAGTTT	66450
Qy	1261	GTGACACGAGACATGTCATAGTCGCCACGGAATCCACTGAACTGTCGACCCAGTTTACC	1320
Db	66451	GCTGATCAGGACATGGGTGTGTTGGACCTTATATCACAGCGCTCCACACGATCACC	66510
Qy	1321	TTTATCAACCAAGTTTGTAAAAACAGCACCTGTCTTTTAAATGAAAAACAAAAAATGC	1380
Db	66511	TTCCGCCATAATTTGTCAAAAACAGCTATGCTACCTTACTGGAAGACCAAAAAATGT	66570
Qy	1381	ATATGTGGTTTTCAGGGTACGTCTTTGAGTTCAAAAAGATTCCTAGACAGCGAACTCTTC	1440
Db	66571	ATATGCGGTTTTCACGGTCTATACAAAGATTCAAGTCCATCTTAGACGGGAGCTATTT	66630
Qy	1441	GTGAGAGCGATTCGCAAGCCGTCGGAATACCTATACAGTTCGTTCTTAAACACGCTACTG	1500
Db	66631	ATCGAAAGTCTATCGCAGATTAACCCGCTTATGTGTACAGTTTCTTTAGTACCTTGTA	66690
Qy	1501	TACACGCGATGTACTGTCTTTCAGCGCTACCGCGTGTAGGTCAACAGAAATATCTG	1560
Db	66691	TATAATGCCATGTACTCATTTTACCGCACGGGTGAAGCAGGGGCATGAAGAATTCTCTC	66750
Qy	1561	CAAGACTAAATTTTGACCCCTCCCGCCGCTCTGGCAACCGGGGTGTAGACCTTCAA	1620
Db	66751	AGGGACCTCAGGGAACTGCGGCTGTCTCAAGAGCTGATCTCTGAGATGAGTCCAGAGAC	66810
Qy	1621	ACGGTTCGTGAAGAGTTAAACCTGGAAGACGACATCTTCTACACGCTGTGAGTCCCCG	1680
Db	66811	GTTCTGGGGCAGAGGGGGACACAGATG-----CCTTCTACCTCACCGCAGCTCCCA	66864
Qy	1681	CCACCCGCGGTATCACTCCCTCAGGTTTGGTCGACAGCTACTGCGCTTAAGGAC	1740
Db	66865	CCATCCCCACCCACCGCGCTCTTCCAACTGTGTGGCTTATCTACCGGGGCCAAGAA	66924
Qy	1741	GTGTTGCGCTCCAGAATAAAGGTGCGGTGTGCTGTTTGGCGGGAGTTTTGAGAAGAA	1800
Db	66925	CTATTCTGCAACAGCTGGCCCTGGCAGCCGACACTTTGTGTGACGAGTCTCTCCACTCC	66984
Qy	1801	ACGTTTTCCGGTTTACGGTTAAATGTCGTAAAGGACGAGTTGACTTTGTCTCCCT	1860
Db	66985	GAATTTTCAACGTTTACCGTGAACATCGTGTGGAGATGGCGTGGACTTTGTGTCCTACT	67044
Qy	1861	TCAGAACCTCTCAACGGGCTGTGGCTTTCGATCGACCGTTGAATCGTATAAAATTAAG	1920
Db	67045	TCCCCCGGGTTCACCGTCTAGTGGCATACGATCCACTATAGACCTATATATCCAG	67104
Qy	1921	GGGTACAGTTTTTACCGGTAGCGTTCGGTTCGCTGTCTAGGGTTTCCCTCAGTGATGAC	1980
Db	67105	GGATATACGTTCTCTCCCACTGAGATTGCGCGTCCAGGAGGACAGCGCTCAGCGAGGAC	67164

QY 1981 CTGAGGAAGATGCCCTCCCTGGTGTGTCAGGACTCTAGCGGTTTATCGCGTGCCTA 2040  
Db  
67165 CTGGCAGAAGATGCCCTCCATAGTGTGTCAGGACTCATCGGGTTCATGCGCTGCCG 67224  
QY 2041 GAGAAATACATAACCAATTGACCGAAACCATGGAGGACGGAGCGCTTTCCAGTGTGC 2100  
Db 67225 GAAATTAACGTCACCAAGATGACAGAGACCTCGAAGGTGGCGACGTGTTTAAACATATGT 67284  
QY 2101 TGTGCGGGGACTATGGGTGTCAGTCAAAATTTAGCCATGACCATCGTAAAGGCACAGGGA 2160  
Db 67285 TGTGACGGGACTACGGTATCAGTTCTAATCTGGCTATGACCATAGTGAAGGCACAGGGG 67344  
QY 2161 ATGTGCTGGAGCGAGTAGCGTAGTATTGGGTCCACAAAGAACCTCCAGACAAGCCAC 2220  
Db 67345 GTTTCACCTAAGTAGGGTGGCCATATCGTTCGGCAACCCAGCCATATCAGAGCCAGTCTA 67404  
QY 2221 GTGTATGTAGCAATATCAAGGGCTGTAACTCAAAATTTATGTCATGACGACCAACCC 2280  
Db 67405 GTGTATGTGGTGTATCCAGGGCCATCGACGCTGCTTACCTGTGTATGACGATATCC 67464  
QY 2281 CTTTAAACCCCTCTCAGAGAACCAGTCGATAACACCTCCGCCAAGCATATAGTCCCGCC 2340  
Db 67465 CTTAAGCTAATGGACCGCGGTGAGCCCGACGTCCCATCTCTCAAAGTACATCAAGGC 67524  
QY 2341 CTCACAACCCAAACACACCTCATCTACTAA----- 2373  
Db 67525 CTATGCAACCCCAAGACTACTCTGTATCTAGTACCTGACCCCGTACCCCTCTCTTAGGACACTGAT 67584  
QY 2374 -----ATAAATAAATAAAGCAAAATGGAACATTTGTG 2406  
Db 67585 GTGTTTGGGAATAAAGCATGAGACTTGACACCTTAAATGTCTGTATTGACACCAATCTT 67644  
QY 2407 TTTTATTCAGTCCAAACAGCGCCACCGTAGTGTCTATTTCCACACACCGGGGGGTGG 2466  
Db 67645 TTTATTTATCAGTCCAGCCACCGCCAGTTATATGACCGTTTCCACACAGGGGTGGCGTGG 67704  
QY 2467 CCGCCATAGTTTGACGACCAAGAACCGCCCGGCCCTCTGGGTTAGAGTAGCCCC----- 2521  
Db 67705 AGGCAGGATCGGGTGTGGGTGCTGACCTGGACCCCGGGTAGTGTCTTCTCTGATG 67764  
QY 2522 -----GCCCGCCGTGGTCGAGTCTATCGTCTCTTAGGGCCCTGGGTT 2564  
Db 67765 AAATCGAGTGGCGGAAAGTACTGGGAGATTGGGTGGGAGGTGACCCCTTTGTCTCGAG 67824  
QY 2565 GTACCAACATAAATCAGTGTCTATCACCCTGTCAGTCCCTGCGTACGCGCGCTCTTG 2624  
Db 67825 GAGACAGATCAGGTCACGGCGACGAGGGCTCTCGTGTGTGCTACTCCCGAGGATA 67884  
QY 2625 CAGGCGGACGAACACCGCCGCTTTGTAAACGTGATCTG-----CAT 2664  
Db 67885 TAATTTACAGACGCCACTGCTTTGGGCTTAAAGTTTGGTTGTCTCTGCGACGCGACCA 67944  
QY 2665 CGTGTGCTGTCTGAGATGAACACAGGTCAATCAATCAGTCGGCGCCGCTGCCCTCAT 2724  
Db 67945 CATCTGCTTACCAGAGAGCGGTAGACTGCCCTTTTGGGCTTTGCGCCCAACGTCATGA 68004  
QY 2725 CGGCTGTGATTCTGTCCCATGTCTCAACAGAGCCTCGGTCCCACTTCGTCTGTCATCTG 2784  
Db 68005 GCCCGATTCTGTACTCAATCTCCCTTGGTCTTCTCCGCTCTCTCGTCCGACGAGGGTG 68064  
QY 2785 AAATAAATAAGTCTCTTGTCTGCGTGTATCCAGGAAACAGAGTTCTTCTCATCGG 2844  
Db 68065 GCTGTGGGAAATAAGCGCGCTGCGTAAACCGGGCTCATTTGTTCACTCGGAGAGT 68124  
QY 2845 AAGACCATCCGCCCTCTCGCGGTGATCTAGATATATACCTGACAGCGGGCGAGGCC 2904  
Db 68125 TGGAACTGTATCGATCCGATGTGAGTCCGATGTGAGTCCAGATCGCGGTGGTGGCGGC 68184  
QY 2905 G----- 2905  
Db 68185 GCAGGGGCGCCAGGAGGCCCTTCAATCAGGCTGCTGTATGTTGAACTTTGTGTTCCAG 68244  
QY 2906 -----TGCGAAGACCGCTCAGCTTCTCCGCCACTTTTCATCAGTGT 2946

68245 GTACACATATTTCTGGAAGCAGGTGAAGTCGGTATGCCCGGTCCCAGTGTATGCCGCCA 68304  
QY 2947 CCTCGTAGTCACTATATAAAAAATCGCTTTCGCTGTCTGATTCAGATGACTGGGGTGGC 3006  
Db 68305 TCGGTTCCAGATAGCAACCCCTCGTCTGCTGAAGGTGAGAGCCAGCGGGAATATC 68364  
QY 3007 GGTGTGGCCAGTAACCAACCGCAACCGGCCCGC-----GACCACCAACCGC 3053  
Db 68365 CGTCATCTGACTTAACCCATCCCATGGACGCTCGGACTCGCGGTGCTGCTGTAACCTGC 68424  
QY 3054 GTCAGTGGCGCGCGCGCTCGGTGGTGTGGTGACCTGTGGAAGTAATGACCGTAGGT 3113  
Db 68425 GCACGCGCCCGCTACCACTGCTACCGGTTGGGGGTATGGGCCGCTGTCGCGAGAGCC 68484  
QY 3114 AGCCGCGCGCTCTGTTAAAGG-----ACTCACGCTCCCGTG 3150  
Db 68485 TCGGCGCAAGTGAGATAAAGGTTGAAAAAGTCTGCGGGTACCCCTCTGGCTGCTT 68544  
QY 3151 CGTGGGATCTATTCTAGGTGGGAGGAGACCGGGTTCGCTTCTTCAATCCACATCAC 3210  
Db 68545 CTTCTCTGAACATCGTCACTTCTTCTCATCTTCTCATCTCTGCTCATATTCAG 68604  
QY 3211 TGTAGTCTGCTCTTCTCCGAATGAGAGTGGTTCGTCTGGCGTTTTCGAGTAGTGGAG 3270  
Db 68605 ATTGCGCGCTCGACTGATCCGGGATATCTGTAGATCCAGAGGGTTCCTGGCGCGATG 68664  
QY 3271 GGGATCTTGTCAAAAAATCCTCAGTGCAGTCACTCAATCACTGGGGGCTCCGT 3330  
Db 68665 CGCTCTCTCGCGGAAGACGCTGCTGGGGCAGACATATCTATCACCGTGGTCCAGCAT 68724  
QY 3331 ACCCGGGCGGTACCGCAACCCGCTACCGGA-----CCCGCGCAACAAAAAC 3378  
Db 68725 AGCCGCGCGCTCGCAATCTCTGGAAGTGAAGAGGTGGAGTGGGAATATGAACT 68784  
QY 3379 GTCTGTTCTCTTTTCTAGTCCCGGGAATCGGAGCATCTCGGAGTGGCGCGC 3438  
Db 68785 TCACGCGGGTCTGTCGGGCGCTCTCTCAATTTGGAAGCATCTCTCTCATCTGTTG 68844  
QY 3439 GGGTCTTTCAGCGGTGCGATAAGAAC-----ATAGCCATGCCC 3476  
Db 68845 TCTAGACGAGTCTCTCAACAACATCGCCATCGGCTTACCGGTTCACCGTAGGGC 68904  
QY 3477 GAAACGTTCACCTGTAACAGCAGTGCCTATCCCGCATGAGTTATATTCAGATCAACGG 3536  
Db 68905 GGAATTTACAAAGCACACGAGTTATTCCTTACTGCTCCACAGGCGCCAGTCCACAG 68964  
QY 3537 CACCCCTCGATTCTGAACATAGATAGTTCGCAAGTTTAAATGGTTGAGCCCTAGAA 3596  
Db 68965 TCTACGCGGTGGCGAGTCAAAATAGTCTGCTAGGTAAAGTGAATACAGCCCTGA 69024  
QY 3597 ACTGCGCGCAGGAGAACCGGTTCGCGTGGCGGTAGCCAGGGGAGACGATGGCGC 3656  
Db 69025 ACCGAGGCCATCGCGAGTTCGGCCACCAAGAGAG-----GCCAGCGGAGATGGATGCTGG 69081  
QY 3657 CTTGAGGACGAGATGCTCTGCTCGGTTTATCAGCGTCCCTTGTCTATAGCTTCTTTC 3716  
Db 69082 CGGTAAAGCACCGGTGTTTCTGTGGTTATGAGCGGAGTCTGTCAATGGCTTGGCC 69141  
QY 3717 CCCACAGCATAAAAACGACCTCTGAAGCTTGGCAGAAAGGCGAGTTTATGTAGTGG 3776  
Db 69142 CCCACAGGAGAAACGCAATGTTCTTAATTTGAGGATATGCTACTGATGATGAACACTCG 69201  
QY 3777 TGAACCAATCCAGCCAGGTTGGGTGTGACCCCGCTTCCCTCTTTCACCGCTCAGAA 3836  
Db 69202 TGAACCAATCCAGCCAGGTTCCCTCGTGTAGCGCGCCCTCTCTCTCCACCGTCAAAA 69261  
QY 3837 TGGTGTTTAAAGAGCAACTCCCGTTTTCGCAACAGTCTTAAGCAACCGCTGAGAGGAG 3896  
Db 69262 CTGTGTTTAGTAGCACACACCTCGCGAGCCAGCTGTCGAGGACCCCGTGGGAAGGAG 69321  
QY 3897 CGGTGAACCCAGGATCAGTATGGCTATCTCTTTAAAAATATTTTGGGCTTGGAGTGA 3956



Db 69322 TACTGAAATGGGACGGAAGCCCTCTAGCTCTCTAAGAATGCTTCTCAAACTGGGTGGAA 69381  
 Qy 3957 TACTGTAGTCCGAGCAACGCTAAAGCTAGTCCGGTGGCTTGACCGCGATGGTAAGGAT 4016  
 Db 69382 CCTGACATTCGGGATCCACACTAAACGCCAGGCCAGTAGCTTGGCCCTTGTGACGGT 69441  
 Qy 4017 CTTGGCCAAAGATCAAGACTTTTAATATCTCTCCGGTTCCGAAACAAATACAGACCCACACATAA 4076  
 Db 69442 CCTGGCCTAAGATCAACCACTTTAATATCTCTGGATCGCAGCAGTGGGACCCACCATCA 69501  
 Qy 4077 TTCTATCTATCGGTGGATTAATCACGGTTAGCTTGTCTATATCCATAACCGCTTCAGCA 4136  
 Db 69502 GCTTGTCTCTGGGGGATACACTGTGGTGTAGCTTAAGTTCCCGAATCTGTCTAGCA 69561  
 Qy 4137 GAGCCGAGCTTCTCTTTTAAAAATGGAAAGATTTTAAAAATTCAGCCATAGTCCG 4196  
 Db 69562 GCGAGACAGCTTCTCTTTTCAAGAAATGATGAGAGGCTCAGAAAGAAATCCACTAGGTG 69621  
 Qy 4197 TAAGTAGCA-----AGTTTGTGTGAGGGTTCTTCAACT 4234  
 Db 69622 CCAAGTAAAGATCCCGGTCCGTCACCCCTGATGATGAGTGGTCCCTTAAGACCG 69681  
 Qy 4235 TCTGGTGACATTTTACTCCAACTATTTGTTTAAAGCAACCCCTCCATAAATCCCGCGTC 4294  
 Db 69682 TCTGTTGCAACCATCGGTCCATGTTGAATTAATTTTCCCTTTGACCTGGTGGCTCTC 69741  
 Qy 4295 TGTGTAATTAATTACGTAATAGTGAACTTTTGTACTCTGTCTGAGAAATCCACAGAG 4354  
 Db 69742 CGGTGCTGCTTTTAGCCCGAGTCTGACTTCCGCTAACAGAACTGTCCGGTTTCATGGCC 69801  
 Qy 4355 GTGCGCGGATTAAGCAGTTTCCGCTTCACTACCGTTTATGTAAAGT-----CT 4404  
 Db 69802 TTTCCACGCTTATATAATTTGTTTACGTTGTGATAGAGCTATCTGCAGTGGTCCGG 69861  
 Qy 4405 TCAAACTCGGTAAAACTCATCAAGACTATATTCATATAATTCACAAATGGAAGCTCT 4464  
 Db 69862 TTAACACCTACAGTATAGGCGGTCAAACTTCGTTGTAATATCCACCAACCACTCAGGTTT 69921  
 Qy 4465 TCTCTGGGACACTGTAAATTTACTGTGTAACAAATCTTATATAAAGCCCTTAAGTGA 4524  
 Db 69922 TCTGCGACGCCAGGACCCCAATCTTCGAACACCGCGACTAAATAATGACCTCAGATTA 69981  
 Qy 4525 AATCCATTTACACATATTTCTGTGTAGAGCGATCAGCTTTATATCGCAATGTTGCTGTA 4584  
 Db 69982 AACCATTCAGCATGTTTCCACGGTAATGTCGCCCTGTTTGTCTGCGAGTTGGCTATA 70041  
 Qy 4585 CACAGCCCTGGACATAAAGCTAAATTTGTTTAATGAAAGAACGAATCTCAACGGCTCC 4644  
 Db 70042 CAGACCCCGTTGCAGTGATTCGATCGCGAAGTGGATAGTGGACCGCAAGAACAA 70101  
 Qy 4645 GGAATATAAAGTCACTAAGTGGTATTTATTTTACACATCGAATTTACTACAAGCCCGGTA 4704  
 Db 70102 GGCAGGTAGAGGCTCCGATGCTGAAATTCGCGACATGTAAGGCGAGTATGGTGA 70161  
 Qy 4705 CAACATGTAATTAATAAATAATCCACATGTAATAATAAATTAAGGTATACAT---ACTT 4761  
 Db 70162 GATGTGACCAATAGGTGGTCCACAGGACGGCAATAGCGCAAAAGATCCCATGGGGCAA 70221  
 Qy 4762 CTCATGTTTTATACATAATATGACGCTAATGCTGTACTTATATAATGATATAGAAATTA 4821  
 Db 70222 ATCCGGGTTTCAACCCTGTGTGCTGGTTCGGTCTCCCGAGGAGCCCTTCGGTAA 70281  
 Qy 4822 TGCCAAATACAGTCTATGTAAGTATATTTTAAAGGCATTAATCCAAACATCAATAATATAAC 4881  
 Db 70282 TATCTGTTTTATATAGTGGGTTTCAAGCATGCGCGAGTCCCGACTAATGAGGACAATTA 70341  
 Qy 4882 ACAAAATGACGCTTCGGATATCCGCCAGCTTTAAACCGAATAATACAGAAATGACTA 4941  
 Db 70342 CTGAAATGACCTTTTCCGACACGGGGGTGAGTCTTATTTCCACGACATATCTCCGG 70401  
 Qy 4942 CACAAACACACCTGAAACCAATTTTATCTCAA-----CATATGCAACAATATT 4990  
 Db 70402 GAAAAATACCACTCTTTAATTTCCGTGGGAGACGATGGGGAATGTGCAATTACC 70461

Qy 4991 TCAGGTAAACCAATGTTTAAATAAGCAATAGCAATGTTTGTAGCTAAGCTCTCTCAA 5050  
 Db 70462 TGACACGGTTCATCAATCATACTCATCGTGGAGCTGTCAACAGCTGTGGCTGAGATTTCTA 70521  
 Qy 5051 AGACAAATTTCTCTCTCAATGATGACTCATCAAC-----CTCGGTTCCTCGTGTAAATGTC 5105  
 Db 70522 AAAAGTCAATCAATGAATCATCGGAATCATCAGACACTCTAAGAACTACTCCATATGCG 70581  
 Qy 5106 GGAATAATATAGTCCAGGATA-CCCTGCTCATCGCCACAAATTAATTTCCAGGGCTTGTGA 5164  
 Db 70582 GGGTGGCGGGGTCCCGAGTAGTGACAGTGGCCATCGGAGACACAGATGATGGGTTG 70641  
 Qy 5165 AAATATCATCCAGTGAATCAGCTGTAATGTCAGAGGATGTTGTTTCTAATTTTAGCCCAT 5224  
 Db 70642 AAATGTCCATACGGGCGGTGTGCAAGGGTCAAGTCCCATCCCAACACAGAGACCTT 70701  
 Qy 5225 CCAATATGTTAGTGGTGACACTTTTGGGACAAAGACCCCGTAAATTTGGCACGCTCCAC 5284  
 Db 70702 TAGATACCTCTCCGGCATGTGCGGTATCCGGGCAAGCAAGCTGTGTCTGGAATCC 70761  
 Qy 5285 AAAAGGAGGATGATTTGATCGTTTACAGAGGGGACGATATTTATTTTATTAACATCAG 5344  
 Db 70762 AAAAGTGGCCAGCGGTACCCAAATCGCAGGGGCTGTATTTATTTTCCACAGAACCG 70821  
 Qy 5345 GAAACGGACCCGTCATGACGACATATGTCATTAAGATGGAGTCTGTCACACTGGTTG 5404  
 Db 70822 GTTCTCTAATTTGATCACACAGGGTATCCAAAGCCGGCTTCCACGTTGATCCGGCTTA 70881  
 Qy 5405 AATTGACGGTGAAAAAATGATTTGCTGTCAATGTCGTAACTATTTTAAACACAGCA 5464  
 Db 70882 CCGACAGTCTTTTCCAGGGTTTCTGTGGGGCGCGGAGCTGACTCAAAAGAGTCACTG 70941  
 Qy 5465 GGGCGGAATAAAACGGGTCCCGCGCTAAGGCGGTGAATCGGCGAGCACAGGCTCAATG 5524  
 Db 70942 CCTTGCCTATGGGCGGTGGTGACAGTCCGCAATCTCTTCCAGGACACTGGCCATGC 71001  
 Qy 5525 CCATTTCCAGATTTGTTAAATCTGTTGAAACATACGAGTCTAAATCAAGTATTAACACG 5584  
 Db 71002 ATGACTCCAAACGCTCTCAGTCCGAGGTAAATGTGCTCTATGAAGATGTGGTAGGCCAGC 71061  
 Qy 5585 ACAAGTTCAACAGAGGCAACGGGTGGTTTCCACCGGACATCCCGCTCACAAACC 5644  
 Db 71062 AGACGTTCAAAACAGATGAATCAAGCTAAGCTCCCGCGGAATCCACATCCCAAGG 71121  
 Qy 5645 CGAACTCTTAGGAACGACATCAAGGTTTCAAGCAAAATTCGAAACCCGAAACCTTAG 5704  
 Db 71122 GGATTTGCTCGGTGCTGTATAGTCTGGAATAGAAACTCAGAAAAAGACACTGACC 71181  
 Qy 5705 TTCTGAGCACCATCTGTGTTATAAAGTGGTTAAGAGAGAGCGGCCACACACTGCTGCT 5764  
 Db 71182 CACCAAGGAGAACCTGGGCTCTTGCAAGTTGATGAGCCCGCAAGAAAGATGTGCTCC 71241  
 Qy 5765 TGTGAGGATTCAGCTTAGTGGCTGCGCTGAGGCTTGCTGACAGGCCCGTAGCTGCGCGG 5824  
 Db 71242 CGTGGACAAAGAGCTTGGGGGGCAGAGATGGCGCTACAGTGGGTGATTTCTTCTACCA 71301  
 Qy 5825 CGCTCGCCGACCTTTGAACGCGGACGCGGATTTCCAGAGCGGAGGAGCGCTTTTTC 5884  
 Db 71302 CGGTCTATCATTTGGTGGACCCACAGGCTCTTCCAGTATCAGCATATAATCTATCTTGC 71361  
 Qy 5885 AGTCCGCGCATATCTCGGCCATAGAGATTCGTCGCTGATCTTTTAGGTGTCATATATA 5944  
 Db 71362 AGTCATCCAGATCAAGATCATGTGATGCTGTGCTTGGCATTTTGGCCGATGTACA 71421  
 Qy 5945 TCTGTTGGGTAAACAACTCACTCCCTCTAAAGGGGAATAGGTCCTGTTTCTATCG 6004  
 Db 71422 TTTCTGTCGCCACATATTTTAACTCTGTAACTAGTAAAGTGAATTCAGTCTGTGTTGA 71481  
 Qy 6005 TTTTCTCTGAATCTATAGTTAGATCTTTTAAAGTATTAACCAAGACCTTAAACCCAGAT 6064  
 Db 71482 GCCCCCGGGGGAACCGAGCTGATGCTTCCAGGACCCACCGGACCTTAGAACCCCGG 71541

72620	ACCAGGTACGGCTCGCTTTGTTCGGTGTCTGHCCTAATATCTGAAATGGTCTTTGCAAGGTAT	72679	TTTAAAGATGGTCTTTCCCGTAAATGCGCACTCATGTCFGRGTGTTTAA-----AC	7352
7184	AGGCCCGTTCCACTCTAAGGGTTGCCGCTTGTGTAGACGGATTCGGGGCAATGTACCGTCT	7243	GTTCACAGGGTGTAAATAACTCAGAAGTCTCCGCTGGACAACTTTTGGCGGACTCTCCG	7412
72680	AGGCTCTTCTCAACGTTTATAGACCGGTACGTGGCAGTCTGTGATTTAGGGTGGCGACGGAC	72739	TCTCTCTTAGGGTAAGAAGCTTCGGCGGTCTCTGTGTGGAAAGCTTCGTCCGCTCTCGG	72919
7244	AGTGTGTCTATTTACCGCACTATCATCCGAGTCGCGATAGTCGGAATCTGTACATATGA	7303	AACCGCTCAGGATAAGCTGTACCACCGGTGCTTCTTATAAACAAGACATACGTCAAAA	7472
72740	AGGTAATCTAACTCTCTGAAGTATCTGATCCAGACGGGTAAATGATACCTTAAACAGATGG	72799	TTGAAACAGGTGATCTTTAAAGGGCGCTTCTCGATGCTCATTTGTAATAAATGACACGCGAC	72859
7304	TTTAAAGATGGTCTTTCCCGTAAATGCGCACTCATGTCFGRGTGTTTAA-----AC	7352	GTTTCACAGGGTGTAAATAACTCAGAAGTCTCCGCTGGACAACTTTTGGCGGACTCTCCG	7412
72800	TTGAAACAGGTGATCTTTAAAGGGCGCTTCTCGATGCTCATTTGTAATAAATGACACGCGAC	72859	TCTCTCTTAGGGTAAGAAGCTTCGGCGGTCTCTGTGTGGAAAGCTTCGTCCGCTCTCGG	72919
7413	AACCGCTCAGGATAAGCTGTACCACCGGTGCTTCTTATAAACAAGACATACGTCAAAA	7472	ACGAACCTGAAGGCCCACTCTACCAGTGTGTGCTCTTATAAATGACACATCAAGAAACA	72979
7473	AACATGACCCACGTGACCTTAATAATGATCACAAGAAGTCTGTGCGATGTATGTAATGG	7532	TCTACGATCCCAAGTACCTTAATAATAGATGTGTGGAAGATGTGTGATTCGGATTTGAAAG	73039
72980	TCTACGATCCCAAGTACCTTAATAATAGATGTGTGGAAGATGTGTGATTCGGATTTGAAAG	73039	AATGTGCTCCTCCGCTGCCCACTCTCCGGGCTCATCGGGATCTGAATCTCTTTAACCTGT	7592
7533	AATGTGCTCCTCCGCTGCCCACTCTCCGGGCTCATCGGGATCTGAATCTCTTTAACCTGT	7592	AATGTTCAAGCTTTGGTCGGCTATGTGTCTGTTACAGACATTAACATGTTTAACCTTT	73099
73040	AATGTTCAAGCTTTGGTCGGCTATGTGTCTGTTACAGACATTAACATGTTTAACCTTT	73099	TCTGCTCTACCGGGGTTCCGGGTTAAACCCCGGGGCGCCACCTGTATACGTCCCGT	7652
7593	TCTGCTCTACCGGGGTTCCGGGTTAAACCCCGGGGCGCCACCTGTATACGTCCCGT	7652	TCTGCTTTTTCGTGCTCTCGAATGAGGACCAAGCGCGCGGTACCAACGTGGCCAT	73159
73100	TCTGCTTTTTCGTGCTCTCGAATGAGGACCAAGCGCGCGGTACCAACGTGGCCAT	73159	GCGCAGATGCGCCACGGGCATCGTGAGAAATCTCACCGAGAGGCCCTTTGTTGTACCG	7712
7653	GCGCAGATGCGCCACGGGCATCGTGAGAAATCTCACCGAGAGGCCCTTTGTTGTACCG	7712	GCGCAGAGGATCCCAAGCAATTAATCGGATCTCTACGAGAGGATCTTATTCGCACAG	73219
73160	GCGCAGAGGATCCCAAGCAATTAATCGGATCTCTACGAGAGGATCTTATTCGCACAG	73219	AAAAGATGTTTCATAGCGTCTCCTCGACGGGGTGTGTATACGCGCCCGAGTTGGCCAGAG	7772
7713	AAAAGATGTTTCATAGCGTCTCCTCGACGGGGTGTGTATACGCGCCCGAGTTGGCCAGAG	7772	AAAAGCATTTCTGACAGCGCATGACGCGGGTGAGCGTCCCTCCAGCCATATGTAAGC	73279
73220	AAAAGCATTTCTGACAGCGCATGACGCGGGTGAGCGTCCCTCCAGCCATATGTAAGC	73279	TTCTGCACGCTCTACCGGAATTAAGGCCAAGTGTTCGGGGCGTGGGGTCTGTCTCA	7832
7773	TTCTGCACGCTCTACCGGAATTAAGGCCAAGTGTTCGGGGCGTGGGGTCTGTCTCA	7832	TACTACAGAAATATACACTGAAATGAAGGCCAAATGCTCGGGGCTCTGGAGGCGACTCG	73339
73280	TACTACAGAAATATACACTGAAATGAAGGCCAAATGCTCGGGGCTCTGGAGGCGACTCG	73339	TATGTTGACGGCGGCTCATATGGCTATGCGGACTCGGTCTCTGCTACTTATATATACC	7892
7833	TATGTTGACGGCGGCTCATATGGCTATGCGGACTCGGTCTCTGCTACTTATATATACC	7892	TCTGCAATCGAGAGGCCATTATGATATTAACCTCTTCCCTACTGAAGCTCTACAAACGT	73399
73340	TCTGCAATCGAGAGGCCATTATGATATTAACCTCTTCCCTACTGAAGCTCTACAAACGT	73399	TGATGCGCAGGGAATACTGAACTTAGACTCAAAAGACTCTGCGAACTGTGTTTTCAAC	7952
7893	TGATGCGCAGGGAATACTGAACTTAGACTCAAAAGACTCTGCGAACTGTGTTTTCAAC	7952	ACGATACCGCGCGGCTCTCTGAGCAGTCCAGGCGCTCTGCTCTTTGGTTTTCCAAAC	73459
73400	ACGATACCGCGCGGCTCTCTGAGCAGTCCAGGCGCTCTGCTCTTTGGTTTTCCAAAC	73459	CGATCTTTTCAAGAATCTTAGCCCCCTATGCACTACTGGCCCAACGGGAGATGGTTC	8012
7953	CGATCTTTTCAAGAATCTTAGCCCCCTATGCACTACTGGCCCAACGGGAGATGGTTC	8012	CGGTCTACTTCCGAGGATTATGGCGCGCTGGAGATCATGACCAAGGCTCAGCTCGCC	73519
73460	CGGTCTACTTCCGAGGATTATGGCGCGCTGGAGATCATGACCAAGGCTCAGCTCGCC	73519	CTGCAAAATATTTTACCATCATCCGGTACGGCGCGAGAGAGGCGCTCTGCTGACTGGAA	8072
8013	CTGCAAAATATTTTACCATCATCCGGTACGGCGCGAGAGAGGCGCTCTGCTGACTGGAA	8072	CTGAAAACTTTTACAGCATCACCGGTTCTGCTGAGAAACGCGCGGCAATATCACCGGCA	73579
73520	CTGAAAACTTTTACAGCATCACCGGTTCTGCTGAGAAACGCGCGGCAATATCACCGGCA	73579	GTACTAGCGGATGACGTGTCCGGAACAGCAGCTGTGCCGATTCCTTAATCTCTGCCAG	8132
8073	GTACTAGCGGATGACGTGTCCGGAACAGCAGCTGTGCCGATTCCTTAATCTCTGCCAG	8132	AGGTCACTGGAATGTCTTATCAGAGAGGGTCTCATGCCAGAACTTTAATTTTGCCAA	73639
73580	AGGTCACTGGAATGTCTTATCAGAGAGGGTCTCATGCCAGAACTTTAATTTTGCCAA	73639	TATGGAACCGGGGCTTCTCCCGGCAACCCCTGGTTGACTCATGTAATGTCTTAGAAATC	8192
8133	TATGGAACCGGGGCTTCTCCCGGCAACCCCTGGTTGACTCATGTAATGTCTTAGAAATC	8192	TCTCGAGCCAGGACTGTGTGCGGCTTCATGTTAGACTCATCAGCGATGTGCTGGCAAAAC	73699
73640	TCTCGAGCCAGGACTGTGTGCGGCTTCATGTTAGACTCATCAGCGATGTGCTGGCAAAAC	73699	CAGAAATCATCTCAGCGGCCACCCCTGAGTCAATTTGTCTATCAAAACACGCAACCCA	8252
8193	CAGAAATCATCTCAGCGGCCACCCCTGAGTCAATTTGTCTATCAAAACACGCAACCCA	8252		

Db 73700 CCGCCGTTATCTGAGCGCCCTGCGCTGAGCGAGTTTGTCTATTAGCAAAACCCCATCCCA 73759  
QY 8253 GTCTGCTCAGTCAGTCAGCAATTATTACCCCAACCCAGGGCGTGTGTTCCCGCCCAATGTT 8312  
Db 73760 ACATGCCGCACACCGTCAGCATCATCCCTTTAAAC--CATCGGTACAGACCCGGCGT 73816  
QY 8313 TTAAGACACCTGGAAAGCGGTGTCTACAGAGCAATTACACACGACGAC 8364  
Db 73817 TTTATTAGTAGTGGCGCGCGGTCTACAGAAATATGTTACAAACATCCAC 73868

RESULT 6  
AR065850  
LOCUS AR065850 35100 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 18 from patent US 5849564.  
ACCESSION AR065850  
VERSION AR065850.1 GI:5996066  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 35100)  
TITLE Chang, Y.; Bohenzky, R.A.; Russo, J.J., Edelman, I.S. and Moore, P.S.  
Polypeptides from Kaposi's sarcoma-associated herpesvirus, DNA  
encoding same and uses thereof  
JOURNAL Patent: US 5849564-A 18 15-DEC-1998;  
FEATURES Location/Qualifiers  
source 1..35100  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 11.0%; Score 1002; DB 6; Length 35100;  
Best Local Similarity 54.5%; Pred. No. 2.1e-253;  
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;  
QY 1 ATGGAGAGTTCCGTCGGATGGACCAACACAGTCGACCAAAATCCGGGTTCATCTTGAAC 60  
Db 29792 ATGGACAGTCGGAAGGGTGCACTGATATGGACGAGCCAGCCCGGGTTTCATCTCAAC 29851  
QY 61 ATGACGTCGATCCCAAGTCAGGGGTGTCTGGATCAGTCAGTCGCTGTCTCAATATA 120  
Db 29852 ATGACATCGATCCCAAGTCAGGAGCGTAGTGGAAACAGATTGACAGAGTTGTCAATATC 29911  
QY 121 ACTACAGGCCACCGGAATGGTGTGTACGACCTGGCCCTCGATCGGCTGAGAGACTCC 180  
Db 29912 ACTACCTGCCCCGAGATGGCTGTGTATGACCTAGAGTTGATCCACTGGAAGACGAA 29971  
QY 181 GGGCCGTTCTGCGTTTACCGTTTATCTAAATTAACGGAACTGTGTGCTGGGAAAGT 240  
Db 29972 GGGCCCTTTCTGCGTTTTCGGCATACGTAATAACCGGGACTGACAGAGCGGGGAAAGC 30031  
QY 241 ACCAGCATATCGGCCCTGTACCAAAATTTAAACTGCTGATCAAGCGGCGGACCCACATA 300  
Db 30032 ACCAGCGTATCCGCCCTTACATCAGAACTCTCACTGCTTAATTACGGGGGCTACAGTGTA 30091  
QY 301 GCGGCACAGAACTATCGGTCGCTTAAAGACGTTCTGTCCACGATCTTCAGCGCTTTT 360  
Db 30092 GCGGCACAGAACTTTCCAGGGCTTTAAAGTCTTACTGTCCCACTATATACACGCTTC 30151  
QY 361 GGGTTTAAGAGCGGACATATATAGCGTCAGAAAGCTCATCAGACCGGAGCGTA 420  
Db 30152 GGATTCAAGAGCAGACACATTAATATCTGCCAGAAAGTGTCCCAAGGTAACTCAGTCC 30211  
QY 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCCAGTATTGGCCGGTTATAGTGGACATT 480  
Db 30212 TCCATCGAGCACTCCAGAGATACGAGCTGGCTAGGTACTGGCCAACTGTCCCGGATA 30271  
QY 481 ATGAAGAGGTTATGGGAAACCAATGGCATGTACGGGACTATATCCACGCGAAT 540  
Db 30272 ATTCGAGAAATTTATGCGCAAGAAACAAAGGGGCGAGTAGTCTCTCTCAAGCGCT 30331  
QY 541 TTTGAAACCTCTCGAGAAATGACGGACCGTGTGTTATGGACTTCCAAATATTATTGTAATC 600

Db 30332 TTCAGACTCTTTTCCGATATGGGTGGAGCCAAATTTGGGACGAGTAACATTATTCGTGATA 30391  
QY 601 GACGAGCCCGAAACCTTGCTCTTTATCATACTCACCACCGTCTGTGTTCTTTTACTGTTTC 660  
Db 30392 GACGAAGCTGGAACCTCTCTGTCCTATTTTGAAGCGGCGTGTGTTCTTCTTATGTTT 30451  
QY 661 CTAAACAGCTGGCTAAATATACCCCTTTTACGCGACGGGGCGGTTCCGTGCATAGTATGC 720  
Db 30452 TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTCCCGTGCCTTGCATAGTCTGC 30511  
QY 721 GTCGGTTTCGCAACGACGACAAACCGCTTCCAGTCAACTTTACAAACACGCGACGCAAAAG 780  
Db 30512 GTGGGGTCTCCACCCAGACGGAGCGCTTTTCAGTCGGTCTTCAACACACGACGACGAGA 30571  
QY 781 ACGGAGATATCGTCTGCGGAGAACATCTTAACATTCTATGATAGGAGAGAGTCTGTATCT 840  
Db 30572 AACGAGATATCTGCTGTGATAATGTCTCACCTTCTTATTTGGGAAAACGTGAGGTTGCA 30631  
QY 841 GAGTACGTACACTTTGGGAGAGAACTGGGCGCTGTTTATAAACAACAAGCGCTGCACCGAT 900  
Db 30632 GATTATATTAGCTGGACGAGAAATGGGCCCTATTATTAACAATAAGCGCTGTACCGAT 30691  
QY 901 CTGAGTTCGGACACCTTCTAAAAATTTTAGAGTATTAATCTTCCCATCTCTGACGAAGTC 960  
Db 30692 CCCAGTTTGGTCACTTGTCTGAGACCTTAGAATATAATCTAGACATATCACCAGAGTTA 30751  
QY 961 ATGAGTTACGTAGACAGATTGTGTTCTTAAAGTAAGATTATGAGTCTCTTTAGATATAC 1020  
Db 30752 ATGACTATATAGTAGTGTGTTGTTCCGAGAGTAGTAAGTTCTGGACCGCTCGAGTAT 30811  
QY 1021 ATTGGCTGGACCCGACTCTTTTGTCAATAGCGAGTAAGGCGTATCTTAACAAACCTA 1080  
Db 30812 GCAGGGTGGACAAGACTCTTCATCTCACACAGGAGGTGAAGTCTTTTCTGGCAACGCTG 30871  
QY 1081 CACACATGTCTAAGCTAGGGGGCGATACGAGGACACGAGGCTCTTTTACTGTCGCCGTG 1140  
Db 30872 CACACCTGCTGTGCGATTAAGGATGCTGTGTCCCAAGGCTTTTACCTGCCAGTG 30931  
QY 1141 GTGTGCGAGGTGTTGTGAAGCCGTTTGAAGATACAAAGCGGCGCTCAACCTCACCAAC 1200  
Db 30932 GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGGGCGTAGGCTCACACAC 30991  
QY 1201 CTCACGCTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTTCCAGTTT 1260  
Db 30992 ATGACTCCCATAGAAATGGGTAAACAAAATCTTTTTCAGGCTTAAGTAACTACTCGAGTT 31051  
QY 1261 GTGACACGAGACATGTCCATAGTCGCCACGGAATCCACTGAACTTCGACCCAGGTTACC 1320  
Db 31052 CCGTATCAGGACATGGCTGTGTGGACCTATATCAAGACGCGTCCACACAGATCACC 31111  
QY 1321 TTTATCAACCAAGTTGTGTTAAAAACAGCCACGCTGTCTTTTAAATGGAATAAACAATAATGC 1380  
Db 31112 TTCGCCACTAAATTTGTCAAAAACAGCTATGCTACCTTACTTGGAAAAGACCAAAAATGT 31171  
QY 1381 ATATGTGGTTTCAGGTTACGTACTTTTGTGTTCAAAAGAAATCCTAGACAGCGAACTCTTC 1440  
Db 31172 ATATGCGGGTTTCAGGGTCTATACCAAGATTCAAGTCCATCTTAGACGGGAGCTATTT 31231  
QY 1441 GTGAGAGCGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTTAAACACGCTACTG 1500  
Db 31232 ATCGAAAGTCATTGCGACGATAACCCCGCTTATGTGTACAGTTTCTTAGTACCGCTGCTA 31291  
QY 1501 TACAACGCGCTACTGTTTTCAGCGTACGCGCTGCTAGGTCTACACGAGAAATATCTG 1560  
Db 31292 TATAATGCATGTACTCATTTTACCGCACGGGGTGAAGCAGGGGCTAGAGAAATCTCTC 31351  
QY 1561 CAAGACCTTAAATTTGACACCCCTCCCGGCCCTCTGGCAACCGGGCGGTGTAGACTTTCAA 1620  
Db 31352 AGGACCTCAGGAACTGCGCGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 31411  
QY 1621 ACGGTTCTGTAAGAGTTAAACCTGGAAGACGACATCTTCTTACCACGCTGTGCGAGTCCCGC 1680

Db 31412 GTTCTGGGGCAGGAG-----GGGGACACAGATGCGCTTCTACCTCAACGCCAGCCTCCCA 31465  
Qy 1681 CCACCCCGGGATATCACTCCCTCCAGGTTTTGGTGCACACGATCTGCGCCCTTAAAGGAC 1740  
Db 31466 CCATCCCCCAACCCAGCGGCTCTTCCAACTAGTGGCCTATTACTCCGGGCGCAAGGAA 31525  
Qy 1741 GTGTTGCGCTCCAGAAATAAAGGTGGGTGCGTGTGTTGGCGGGAGTTTGAGAGGAA 1800  
Db 31526 CTATTCTGCAACAGGCTGGCCCTGGACGCCGACATTGGTGACAGTTCTCCACTCC 31585  
Qy 1801 ACCTTTTCCCGTTTACGGTTAAACATGGTTCGTAAAGGACGGAGTTGCACTTTGTCCTCC 1860  
Db 31586 GATTTTTCACAGTTTACGGTGAACATCGTGGTGGAGATGGCGTGACTTTGTGCTCACT 31645  
Qy 1861 TCAGAACGTCTCAACGGGCTGGGTTTGGCTTGCATCGACCGTTGAATCGTATAAAATTAG 1920  
Db 31646 TCCCCCGGGTCCACGGTCTAGTGGGACGACGATCCACTATAGACACCTATATAATCCAG 31705  
Qy 1921 GGGTACACGTTTACCGGTAGCGTTTGGTTCGTGTCAGGGTTTGCACCTCAAGTGAAGAC 1980  
Db 31706 GGATATACGTTCTCCAGTGAGATTTCGGCGCTCCAGGAGGACAGCGCCTCAGCGAGGAC 31765  
Qy 1981 CTCAGGAGAGATGCCCTCCCTGGTGGTGCAGGACTTAGCGGTTTATTCGGTGCCTA 2040  
Db 31766 CTGGCGAAGATGCCCTCCATAGTTTCCAGACTCATCGGGGTTCAATTGGCTGCCCTG 31825  
Qy 2041 GAGAAATACATAACCAAATTGACCGAAACCATGGAGACGGGACGCTTTTCCAAAGTGTGC 2100  
Db 31826 GAAAAATAACGTCAACAGATGACAGAGACCTCGAAGTGGCGAGCTGTTTAACTATGCT 31885  
Qy 2101 TGTGCGGGGACTATGGGTGAGCTCAATTTAGCCATGACCATCTGTAAGGACAGGGA 2160  
Db 31886 TGTGACGGGACTACCGGTATACGTTCTAATCTGGCTATGACCATGTAAGGACAGGGG 31945  
Qy 2161 ATGTCGTGGAGCAGTAGCGGTAGTATTGGGTGCCACAAAGACGTCCAGACAAAGCCAC 2220  
Db 31946 GTTTCATAGTAGGTGGCCATATCGTTCGGCAACACCGCGAATATCAGAGCCACTTA 32005  
Qy 2221 GTGTATGTAGCAATATCAAGGGCTGTAACTCAAAATTTTGTGTCATGACAGCAACCC 2280  
Db 32006 GTGTATGTGGTGTATCAGGGCCATCGACGCTCGTTACCTGTTAATGGACGTAATCC 32065  
Qy 2281 CTTAAACCCCTCTCAGAACACAGTCGATAACACCTCCGCCAGGATATAGTCCGCGG 2340  
Db 32066 CTTAAGCTAATGGACCGCGGTGACGCCAGTCCCATCTCAAAGTACATCATCAAAGCC 32125  
Qy 2341 CTCACAAACCAACACACCCCTCATCTACTAA----- 2373  
Db 32126 CTATGCAACCCACAGACTACTCTGATCTACTGACCCGCTACCCCTCTCTTAGGACACTGAT 32185  
Qy 2374 -----ATAAATAAATAAAGCAATGAGCAATGGAACATTTG 2406  
Db 32186 GTGTTTGGGAATAAAGCATGAGACTTGACACCTATATAATGTTCTGTATTGACACCATCTT 32245  
Qy 2407 TTTTATTTCAGTCCAAACACGGCCACGGATAGTTGTCAATTTCCACACACCGGGGGGTGG 2466  
Db 32246 TTAATTTACGTCCAGCCACGGCCAGTTATGACCGTTTCCACACAGGGTGGGGTGG 32305  
Qy 2467 CCGCCATAGTTTACGACCAAGAACCGCCGGGCGCTCTGGGTTAGAAGTAGCCCC----- 2521  
Db 32306 AGGCCAGGATGCGGGTGGGTGCGTGCACCTGGACCCCGGGTAGTTGTGCTTCTGATG 32365  
Qy 2522 -----GCCCGCGTGGTGGAGTGCATGCTCTTTAGGGCCCTGGGTT 2564  
Db 32366 AAATCGAGTGGCGGAGTACTGGAGATTGGGTGGGAGTGACCTTTGTGCTCGACG 32425  
Qy 2565 GTACCAACATAAATCAGTGTGTCATACCGTTGAGTCCCTGCGTACGCGCGCTCTTTG 2624  
Db 32426 GAGACACGATCAGCTCAGCGGACGAGGGCTCCTGCTGTGTGTCACCTCCCGAGGATA 32485  
Qy 2625 CAGCGGCACCAACACCGCGCTTGTAAACGTGACT-----GCAT 2664  
Db 32486 TAAATTACGAGCGCCACTGCTTTGGGCTTAAAGTTGTTGTTGTTCTCTGCGACGGCACCA 32545

Qy 2665 CGTGTGCTGTCTGAAGATGAACACAGGTCAATCATCATGCTGGCGCCGCTGCTCAT 2724  
Db 32546 CATCTCGCTACCAAGAGAGCGGTAGACTGCTTTTGGCTTTTGGCCCCACGTCCATGA 32605  
Qy 2725 CGGGTCTGATTCGTCTCCATGTCTCAACAGAGCCTCGGTCTGTCACACTTCGTCGTCATCTG 2784  
Db 32606 GCCCGATTCTCTGACTCAATACCTTCCCTTGGTCTTCTCGGTCTCTCTCGACAGGGTG 32665  
Qy 2785 AAATAAATAATGCTCTTTGCTCTGCTCAGTTGATTCAGCAACACAGGTTCTTCTTCATCGG 2844  
Db 32666 GCTGGTGGAAAAATGGCGCGCTCGGTAAACGGCGCTCATTTGTTCACTCCGAGAGT 32725  
Qy 2845 AAGACCATCCCGCTCTGCGCGGTGATCTAGATATATCACCTGACAGCGCGCGGAGGCC 2904  
Db 32726 TGGAACTGTCATCGCTATCAGAGTCCGATGTCAGTGCAGCATCGCGGTGGTGGCGGC 32785  
Qy 2905 Q----- 2905  
Db 32786 CAGGGGGCGCCACAGAGGCGCTTTCATCAGGGTCCGTGTATGTAAGTCTTTGTTGTTCCAG 32845  
Qy 2906 -----TGGCGAAAAGCCGTGAGCTTCTCCGCCACTTTTCATCAGTGT 2946  
Db 32846 GTACACTATTCTGGAAGCAGTGAAGTCCGTATGCCCGGTCCAGTGTATCCGCCA 32905  
Qy 2947 CCGTGTAGTCACTATATAAAAAATCGTCTGCTGTCTGATTCAGATGACTGGGGTGGC 3006  
Db 32906 TCGGTTCCAGGATAGCAACCCCTCGTCTGCTGAAGGTGAGAGCCCGACAGGGGAAATC 32965  
Qy 3007 GGTGTGGCCAGTAAACACCGCAACGGGC-----CGCGGACCAACCAACCGC 3053  
Db 32966 GTCATCTGACTAACCCATCCATGAGCGCTCGACTCCCGCTGTCCTTGAACATGC 33025  
Qy 3054 GTCAAGTCCGCGGGCGGCTCGGTGTGGTGACGCTGTGCAAGTAAATGACGGTAGGT 3113  
Db 33026 GCACGCGCGCGCTACCACTGCTACCGGTTTGGGCGTATGGGCGCTGTGCCAGAGGCC 33085  
Qy 3114 AGCGCGCGCTCTGTTAAAGG-----ACTACCGTCCCGTG 3150  
Db 33086 TCGGCGCAAGTGAGATAAAGGTTGAAAAAGTCTGCAAGGTACCCCTCTGGCTGCTTT 33145  
Qy 3151 CGTGGGATCTATTCTAGGTGGGAGGAGACCGCGGTTCTGTTCTTTCATCCACATCAC 3210  
Db 33146 CCTCTGAACATCGTCACTTTCTTCTTCATCTTCTCATCTCTCGTCATATTCAG 33205  
Qy 3211 TGTAGTCTGCTTCTCGAATGAGAGTCGGTTTCTGTCGGGCTTTTGGAGTAGTGGAG 3270  
Db 33206 ATTCCGCGCTCGACTGATCCGGGATATCTGTAGATCCAGAGGGGTTGCTGGCGCGATG 33265  
Qy 3271 GGGAACTTGTCAAAAAACATCCTCAGGTGAGTCATGTCAATCACTGGGGGCTCCGT 3330  
Db 33266 GCGTGTCTCGCGAAGACGTCGTCTGGGCGAGACATATCTATCACCGTGGTCCAGCAT 33325  
Qy 3331 ACCCGGGCGGTACCGGCAACCCCGTACGGGA-----CCCGCGCAACAAANAC 3378  
Db 33326 AGCCCGCGGCTGCAAAATCTGGAAGTGAAGAGAGGTGAGGTGGGAATATGAAT 33385  
Qy 3379 GTCTGTTCTCTTTTCTAGGTGCGCGGAATCGGCAGCATCTCTGGAGTGGCGCGC 3438  
Db 33386 TCACGGGGGTGCTGTCGAGCGCTCTTCAATTGGAGCATCTCTCTTCATCTGCTG 33445  
Qy 3439 GGGTCTTTGACGGGTGCGATAAGAAC-----ATAGCCATGCGC 3476  
Db 33446 TGCTAGACGAGTCTCTCAAAACATCGCCATGCGCTTGTAGCGGGTTGACCGCTAGGGGC 33505  
Qy 3477 GAAAGTTTCACTGTAAAAACGACATGCGCTATCCCGATCAGTTAATATTCAGTCAACGG 3536  
Db 33506 GGAATAATTACAAAGCACAGAGTTATTCCTTTTACTGCTCCACAGGCCCCAGTCCACAG 33565  
Qy 3537 CACCCCTCCGATTTCTGAAGTACTAGATGATTCGCAAGTTTAAATGGTTGAGCTAGAA 3596  
Db 33566 TCTACGCGCGGTGGCGAGTCAAAATAGTCTGGCTAGTTAAAGTGAATTACAGCCCTGGA 33625

QY 3597 ACTGCGGCGGAGAACCGGTTGCGCGTGGGGTAGCCAGCGGAGAGCGGATGGCGG 3656  
DB 33626 ACCGAGGCGATCGCGAGTGTGGGCCAACAGAGAG---GCCACGGAGATGGATGCTGGG 33682  
QY 3657 CTTGAGGACGAGATGCTCTGTCGTTTATCAGCGTCGCTTGTCTATAGCCTTCTTC 3716  
DB 33683 CCGTAAGCACCGAGGTTCTGTCGTTTATGAGCGAGTCTGTCAATGGCCTTGGCC 33742  
QY 3717 CCCACAGCATAAAAACGACCTCTGAAGCTTTGGCAGAAAGCGAGCTTATATGATAGCTGG 3776  
DB 33743 CCCACAGGAGAAAAACGCAATGTTCTTAATCTTTAGGATATGCTACTGATGAACCTCG 33802  
QY 3777 TGAACCAATCCCGACGACGAGTTGGCGGTGACCGCGCTTCCCTTTCACCGCTCAGAA 3836  
DB 33803 TGAACCAATCCCGACGACGAGTCCCTGCTGAGCGCGCCCTCCCTTCTCCACCGTCAAAA 33862  
QY 3837 TGGTGTAAAAAGCAGAACTCCCGCTTTTGGCCACAGCTCTAAGCACCCTGAGAAAGGAG 3896  
DB 33863 CTGTGTTTAGTAGCAACACACCTTGGCGAGCCAGCTGTGAGGACCCGCTGGGAGGAG 33922  
QY 3897 CGGTGAACCCAGGTACAGTATGCTATCTCTTTAAATAATATTTTGGGCTTTGGAGSTA 3956  
DB 33923 TACTGAAATTTGGGACGGAAGCTCTAGCTCTCTAAAGATGCTTCTCAAACTGGGTGGA 33982  
QY 3957 TACTGATGCGGAGCAACGCTAAAGGCTAGTCCGCTGGCTTGACCGCATGTTAAGGAT 4016  
DB 33983 CCGTACATTTGGGATCCACACTAAACGCGAGCCAGTAGCTTGGCCCTTGTGCTACGGGT 34042  
QY 4017 CCGTCCCAAGATCAGGCTTATATCTCCGCTTCCGACATAGCAGTACGACCAACATAA 4076  
DB 34043 CCGTCCCTTAGATCAGCACTTATATCTCTGATCGGACGAGTGGGACCAACATCA 34102  
QY 4077 TCTATCTATCGTGGATTAATCAGCTAGTTCCTATCATATCAATCAACGCTTCAGCA 4136  
DB 34103 GCTGTCTGTGGGGATACACTGTGCTGTAGCTTCCGAACTGTCTGAGCA 34162  
QY 4137 GAGCGCGAGCTCTGTTTAAATAATGAGAGATTTAAATTTCAAGCA 4188  
DB 34163 GCGAGAGCGTTCGTTTCAAAAATGATGAGGCTCAGAAAGAAATCCA 34214

RESULT 7  
AR127848  
LOCUS AR127848 35100 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 18 from patent US 6183751.  
ACCESSION AR127848  
VERSION AR127848.1 GI:14115510  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 35100)  
AUTHORS Chang,Y., Bohenzky,R.A., Russo,J.J., Edelman,I.S. and Moore,P.S.  
TITLE Unique associated Kaposi's Sarcoma virus sequences and uses thereof  
JOURNAL Patent: US 6183751-A 18 06-FEB-2001;  
FEATURES  
source  
1. .35100  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 11.0%; Score 1002; DB 6; Length 35100;  
Best Local Similarity 54.5%; Pred. No. 2.1e-253;  
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;

QY 1 ATGAGAGTTCGTCGATGACCAACACAGCTCGAACCAATCCGGGGTTTCATCTTGAAC 60  
DB 29792 ATGACAGCTCGGAAGGTCATGATATGACAGCGCCGACGCCCGGGTTTCATCTCAAC 29851  
QY 61 ATGACGTCGATGCAAAAGTCAGGGGTGTCGTGATCAAGTCAGTCAGTCGCTGCAATATA 120  
DB 29852 ATGACATGGATGCAAAAGTCAGGAGGTAGTGGAAACAGATTGACAGGTTGTCAATATC 29911

QY 121 ACTACGACCGCCACCGGAAATGGGTGTTGATGACGCTGGCCTTCGATCCGGCTGAGACTCC 180  
DB 29912 ACTACCTCGCCCGGAGATGGGCTGATATGACTAGATTCGATCCACTGGAAGACGAA 29971  
QY 181 GGGCGGTTCTTGGCGTTTACCGTTTATCTAATTACGGGAACCTGCTGCTGCTGGGAAAGT 240  
DB 29972 GGGCGGTTCTTGGCGTTTACCGTTTATCTAATTACGGGAACCTGCTGCTGCTGGGAAAGC 30031  
QY 241 ACCAGCATATCGGCGCTGTACCAAAATTTAAATCTGCTGATCAGCGGCGGACCAACATA 300  
DB 30032 ACCAGGATATCGGCGCTGTACATCAGATCTCACTGCTAATTACGGGGGCTACAGTGGTA 30091  
QY 301 GCGGACAGAAACCTATCGGCTGCGCTAAAGACGTTCTGCTCCACGATCTTACGCGCTTTT 360  
DB 30092 GCGGACAGAAATCTTCCAGGGCTTTTAAAGTCTTACTGTCTCCACTATATACCAAGCGCTTC 30151  
QY 361 GCGTTTAAAGACGACACATCAATATAGCGCTCAGAAAGCTCATCAGACCGGAGCGGTA 420  
DB 30152 GGAATCAAGAGCAGACATTTATCTGCCAGGGAAGTGGCCAAAGTAACTCAGTCC 30211  
QY 421 TCCATAGAGCAATTCAGCAACAGGAGCTATCGAAGTATTTGGCGGGTTATAGTGGACATT 480  
DB 30212 TCCATCGAGCAACTCCAGAGATACGAGCTGGCTAGTCTGGCCAACTGTCAACGATATT 30271  
QY 481 ATGAAAGAGGTTATGGCGAAAAAACCCAAATGGCATGTACGGGACTATATCCAAACGGGAAT 540  
DB 30272 ATTGAGGATTTATGCGCAGAAACAAAGGGGAGTATAGCTCCCTCTCTCAAGCGCT 30331  
QY 541 TTTGAAACCTCTCGAAGATGACCGGACCGTGTATTGAGACTTCCAAATATTATGTAATC 600  
DB 30332 TTCAGACTCTCTTGGCGTATGAGTGGAGCCAAATTTGTGGACGAGTAACATTTATCGTGATA 30391  
QY 601 GACGAGCGCGGAAACCCCTGCTCTTACATACTCACCAACCGTGTCTTCTTTACTCGGTC 660  
DB 30392 GACGAGCTGGAACCCCTCTGCTCCCATATTTTGAAGCGCGTGTGTTCTTCTATTGTTT 30451  
QY 661 CTAAACAGCTGGCTTAATACCCCTCTTACCGCAGCGGGCGGTTCCGTGCTAGTATGC 720  
DB 30452 TACAACAGTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTGTCTGCTAGTCTGC 30511  
QY 721 GTCGGTTTCGCAACGAGACAAACCGGTTCCAGTCAACTTACAACCAACGGGACGCAAAAG 780  
DB 30512 GTGGGCTCTCCACCCAGAGCGGCTTTCAGTCCGCTCTTCAACCAACGAGCAGAGA 30571  
QY 781 ACGGAGATATGCTGTCGAGAGACATCTTAAATTTTAGAGTATAATCTTCCCATCTCTGACGAGTCT 840  
DB 30572 AACGAGATATCTGCTGTGATAATGTGCTCACCTTCTATTGGGAAACGCTGAGGTTGCA 30631  
QY 841 GAGTACGTACACTTTGGAGAGAAACTGGGCGCTGTTTATAAAACAAAGCGCTGCACCGAT 900  
DB 30632 GATTAATTAGGCTGGAGAGAAATGGGCCCTATTATATAAATAAGCGCTGTACGGAT 30691  
QY 901 CTGCAAGTTCGACACCTCTTAAATAATTTTAGAGTATAATCTTCCCATCTCTGACGAGTCT 960  
DB 30692 CCCAGGTTGGTCACTTGTGTAAGACCTTTAGAAATATAATCTAGACATATCACCAAGTTA 30751  
QY 961 ATGAGTTACGTAGACAGATTTGCTGTTCTCTTAAAGTAAGATTATGATCTTTTAGAATAC 1020  
DB 30752 ATGAGCTATATAGTAGTTTGTGTTCCGAGAGTAAGATTCTGGAACCGCTCGAGTAT 30811  
QY 1021 ATTGGCTGGACCGGCTCTTTTGTGCTACATAGGAGTAAAGCGTATCTAACAACCTTA 1080  
DB 30812 GCAGGGTGGACAAAGCTCTTCACTCACACCGAGGAGTGAAGTCTTTTCTGCAAGCGCTG 30871  
QY 1081 CACACATGCTAAACGCTAGGGGCGGATACCGAGGACACGAAGCTCTTTTACCTGTCTCCGCTG 1140  
DB 30872 CACACCTGCTGTGAGTAATAAGGATGCTGTGCCAATAAGCTTTTACCTGCCAGTG 30931  
QY 1141 GTGTGCGAGGTGTTTGTGAAGCGTTTGAAGAAATACAAGGGGCGCTCAACTCACCACAC 1200  
DB 30932 GTCTGTGAGGTGTTTACAGAGCCATTTGAGAGTACAAACGGGCGGTAGGCGCTCACAC 30991  
QY 1201 CTCACCGTGCAGGCTGGGTGACAAAAACCTCTTTTAAAGTTAGTAATTTATTCGAGTTT 1260





QY 3211 TGAGTCGCTGTTCTCCGAATGAGAGTCGGTTTCGTCGCGCTTTTGGAGTAGTGGAG 3270  
 Db 33206 ATTGCGCGCTGACTGATCGGGGATATCTGTAGATCCAGAGGGTGTCTGCGCGGATG 33265  
 QY 3271 GGAATCTTTGGTCAAAAAATCTCAGGTGCAATCATGTCATCAATCAATCGGGGCTCCGT 3330  
 Db 33266 GCGTGTCTCTCGGGAAGAGCTGTCTGCGGAGACATATCTATCACCCTGGGTCCAGCAT 33325  
 QY 3331 ACCCGGCGGTACCGGCAACCCGATACGGGA-----CCCGGGACAAACAAAC 3378  
 Db 33326 AGCGCGCGGCTGCGCAATCTCTGGAAGTGATGAAGAGGTGGAGTGGGAATATGAAT 33385  
 QY 3379 GTCTTGTCTCTTTTCTAGTGTCCCGGGAATCGGCAGCATCTCTGGAGTGGCGGCG 3438  
 Db 33386 TCACGGGGGTCTGTCTGCGAGGCGCTCTCAATTGGAAGCATCTCTTCTCATCTGTG 34445  
 QY 3439 GGTCTCTTGAAGGTGCGATGAAGC-----ATAGCCATGGCC 3476  
 Db 33446 TGCTAGACGAGTCTCTCAAAACATCGCCATGGCCTTTGACGGGGTGTAGCCCTAGGGG 33505  
 QY 3477 GAAACGTTACCTGTAAACCGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGG 3536  
 Db 33506 GGAATTTCAAGACACACGAGTTATGCTTTACTGCTCCACAGGCCCTCAGTCCACAG 33565  
 QY 3537 CACCCCTCCGATTTCAACTAGATGATCTGCAAGTTTAAAGTTTAAAGTTTGAAGCTAGAA 3596  
 Db 33566 TCTCACGCCGCTGGCGAGTCAAAATAGTGTGCTAGGTAAAGTGAATACACGCCCTGGA 33625  
 QY 3597 ACTCGGCGGAGGAGAACCGGTTGCGCGGTGGCGGTAGCGAGGAGAGCATGCGCG 3656  
 Db 33626 ACCGAGGCCATCGGAGTGTGGGCCAACAGAGAG--GCCAGCGAGATGATGCTGGG 33682  
 QY 3657 CTTTGAGGACGAGATGTCTCTGCTGCTTTATCAGCGTCCCTTGTCTATAGCCTTTCTTC 3716  
 Db 33683 CCGTAAGCACCAAGTGTCTCTGCTGCTTTATGAGCGAGTTCTGTCAATGGCCTTGGCC 33742  
 QY 3717 CCCACGACATAAAACGACCTCTGAGCTTGGCAGAAAGGAGCTTATATGCTAGCTGG 3776  
 Db 33743 CCCACAGGAAAGACGAAATGTTCTAACTTTGAGGATATGCTACTGATGAAGAACTCG 33802  
 QY 3777 TGAACCAATCCAGGCAAGTGTGGCGTGTGACCCCGCCTTCCCTTTCCACCGTCAGAA 3836  
 Db 33803 TGAACCAATCCAGGCAAGTGTGCTGTGAGCGCGCCCTCCCTTTCTCCACGTCAAA 3862  
 QY 3837 TGTGTTTAAAGCAGAACTCCCGCTTTGGCCCAACAGTCTAAGACCCGTCGAGAAGGAG 3896  
 Db 33863 CTGTGTTTAGTACCAACACACCCCTGGCGAGCCAGCTGTGAGGCAACCCGTGGAGGAG 33922  
 QY 3897 CGGTGAACCCAGGTACAGTATGGCTATCTCTTTAAATAATTTTGGAGCTTGGAGGTA 3956  
 Db 33923 TACTGAATTTGGGACGGAAGCTCTAGCTCTCTAAGATGCTTCTCAACTGGTGGAA 33982  
 QY 3957 TACTGTAGTCCGAGCAACGCTAAAGCTAGTCCGCTGGTGTGACCGGATGGTAAGGAT 4016  
 Db 33983 CCGTGAATTCGGATCCACACTAAACGCCAGGCGAGTGTGGCCCTTGGCTGTGGTGGT 34042  
 QY 4017 CCGTGGCAAGATCAGACTTTAATATCTCCCGTTGCGCAACATACGACCCACCATAA 4076  
 Db 34043 CCGTGGCCTAAGATCAGACTTTAATATCTCTGATCGGAGAGTGGGACCCACCATCA 34102  
 QY 4077 TTCTATCTATCGGTGGATAAATCAGGTTACGTTGTCTCATATCCATTAACCGCTTCAGCA 4136  
 Db 34103 GCTTGTCTCTGGGGGATACACTGTGTGTGTAGCTTAAGTTCGCCAATCTGTCTGAGCA 34162  
 QY 4137 GAGCGCGAGCTTCTGTTTAAATTTGGAAGAAAGATTTAAATTTCAAGCCA 4188  
 Db 34163 GCGAGACAGTTTCTGTTTTCAGAAATGATGAGAGGCTCAGAAAGGAATCCA 34214

RESULT 8  
 AR194750 35100 bp DNA linear PAT 20-APR-2002  
 LOCUS

DEFINITION Sequence 18 from patent US 6348586.  
 ACCESSION AR194750  
 VERSION AR194750.1 GI:20241342  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 35100)  
 AUTHORS Chang, Y., Bohenzky, R.A., Russo, J.J., Edelman, I.S. and Moore, P.S.  
 TITLE Unique associated Kaposi's sarcoma virus sequences and uses thereof  
 JOURNAL Patent: US 6348586-A 18 19-FEB-2002;  
 FEATURES  
 location/Qualifiers  
 source 1..35100  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

Query Match 11.0%; Score 1002; DB 6; Length 35100;  
 Best Local Similarity 54.5%; Pred. No. 2.1e-253;  
 Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;

QY 1 ATGAGAGAGTTCGCTCGGATGACCAACACAGTCGTGCAACCAATCCGGGTTTCATCTTGAAC 60  
 Db 29792 ATGACAGCTCGGAAGGTTGCACTGATATGACGAGCCAGCCCGGTTTCATCTCAAC 29851  
 QY 61 ATGACGTCGATGCCAAAGTCAGGGTGTGTGTGATCAGTCAGTCGCTGCTCAATATA 120  
 Db 29852 ATGACATCGGATGCCAAAGTCAGGAGCTAGTGAAACAGATTGACAGGTTTCAATATC 29911  
 QY 121 ACTACACGCCACCGGAAATGGGTTGTGACACCTGCGCTTCGATCCGGCTGAAAGACTCC 180  
 Db 29912 ACTACCTCGCCCGGAGATGGGCTGTATGACCTAGAGTTCCACTTGAAGACGAA 29971  
 QY 181 GGGCGGTTCTTTCGCTTTACCGTTTATCTAATACGGGAACCTGCTGTGTGGGAAAGT 240  
 Db 29972 GGGCCCTTCTGCGCTTTTCGGCATAGTAATAACGGGAGCTGACGAGCGGGGAAAGC 30031  
 QY 241 ACCAGCATATCGGCGCTGTACCAAAATTTAAACCTGCTATCAGCGGCGGACCAACATA 300  
 Db 30032 ACCAGGTATCGGCGCTTACATCAGAACTCTCACTGCTTAAATAGCGGGCTACAGTGTGA 30091  
 QY 301 GCCGACAGAACCTATCGCTCGCTAAAGACGTTCTGTCCACGATCTTCAGGCGCTTTT 360  
 Db 30092 GCGGCACAGAACTTTCCAGGGCTTTAAAGTCTTACTGTCCACTATATACAGCGCTTC 30151  
 QY 361 GCCTTTAAGAGCCGACATCAATATAGCCGTGAGAAAGCTCATCAGACGGAGCGGTA 420  
 Db 30152 GGAATTCAGAGACACACATTAATATCTGCCAGAGGAAAGTCCCAAGGTAACTCAGTCC 30211  
 QY 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGCTTATAGTGGACATT 480  
 Db 30212 TCCATCGAGCAACTCCAGAGATACGAGCTGGCTAGGTACTTGGCCAACTGTCCCGATATT 30271  
 QY 481 ATGAAGAGGTTATGGCGAAACCAACCCATGGCATGTACGGGACTATATATCCACGCGAAT 540  
 Db 30272 ATTCGAAATTTATGCGCAAGAAACAAAGGGGCGAGTATAGTCCCTCTCTCAAGCGCT 30331  
 QY 541 TTGAAACCCCTCTCGAAGTACCGGACCGGTTTATGAGCTTCCAAATATTATTGTATATC 600  
 Db 30332 TTCAGACTCTTTCGCGTATGGGTGGAGCCAAATTTGGACAGATTAATATCTGTGATA 30391  
 QY 601 GACGAGCGCGGAAACCTGTCTCTTATCATCTACCAACCGTCGCTGTTCTTTTACTGTTTC 660  
 Db 30392 GACGAAGCTGGAACCCCTCTGCTCCCATATTTTGAAGCGGCTGTTCTTCTATTGGTTT 30451  
 QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTTACCGCAGGGGGGGTTCGCTGCATAGTAGTC 720  
 Db 30452 TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTTGTGATAGTCTGC 30511  
 QY 721 GTCGGTTCCGCAACGACAGCAAAACCGGTTCCAGTCAACTTACAAACGAGGACGCAAAAG 780  
 Db 30512 GTGGGCTCTCCACCCAGACGACGCTTTTCAAGTCGGTCTTCAACCAACGACGACGAGAG 30571

Qy 781 ACGGAGATATCGTGTGCGAGAACATCTTAACATTTATGATAGGGAAGAGTCTGATCT 840  
 Db 30572 AACGAGATATCTGCTGTGATATAATGTCTCACCTTCTATTGGGAAAACGTGAGGTGCA 30631  
 Qy 841 GAGTACGTACACTTGGAGAGAACTGGCGCTCTTTATTAACACACAGCGCTGCACCAT 900  
 Db 30632 GATTATATTAGGCTGGACAGAAATGGGCGCTTATTATTAACAAATAGCGCTGTACGAT 30691  
 Qy 901 CTGAGTTCGAGACACACTTCTAAAATTTTATAGATATAATCTTCCCATTTCTGACGAAGTC 960  
 Db 30692 CCCAGTTTGGTCTGCTGAAGACCTTAGAATATAATCTAGACATATACACAGAGTTA 30751  
 Qy 961 ATGAGTTACGTAGACAGATTTGTGCTTCTTAAAGTAAGATTAATGATCTTTAGATATAC 1020  
 Db 30752 ATGAGCTATATAGATAGTTTGTGTTTCCGAAGAGTAAGATTTCTGACCGCTGCGAGTAT 30811  
 Qy 1021 ATTGGCTGGACCCGACTCTTTTGTACATAGCGAGGTAAAGCGGTATCTAAACAACTTA 1080  
 Db 30812 CGAGGTGAGACAGACTCTTCACTCACACGAGAGTGAAGTCTTTCTGGAACGCTG 30871  
 Qy 1081 CACATATCTAAGCTTAGGGGGGATACACAGGGAACAGAGTCTTTTACCTGTCCTCGTG 1140  
 Db 30872 CACACCTGCTGTGAGTAATAGGATGCTGTGTCCCAAAAGCTTTTTCACCTGCCAGTG 30931  
 Qy 1141 GTGTGCCAGGTGTTTGTGAAGCGTTTGAAGGAATACAAAGGGCGCTCAACCTCAACAA 1200  
 Db 30932 GTCTGTGAGGTGTTTACAGAGCCATTTGAGAGATACAAAGCGCGGTAGGCTTCAACAC 30991  
 Qy 1201 CTCACCGTACCGAGTGGGTGACAAAACCTCTTTAAAGTTAAGTAAATTTATTCGCAAGTT 1260  
 Db 30992 ATGACTCCCATAGAAATGGGTAAACAAAATCTTTTTCAGGCTAAGTAATCTACCTGCA 31051  
 Qy 1261 GTGACACGACATGTCATAGTTCGCAACGGAATCCACTGAAGTTTCGACCGAGTTTACC 1320  
 Db 31052 GCTGATCAGGACATGGTGTGTTGGGACCTATATACACAGCGGTCCACAGATATACC 31111  
 Qy 1321 TTTATACCAAGTTTGTAAAACAGCCACGCTGTCTTTAAATGAAAACAAAATAATGC 1380  
 Db 31112 TTGCGCACTAAATTTGTCAAAAACAGCTATGCTACCTTACTGAAAAGACCAAAAATGT 31171  
 Qy 1381 ATATGCGGTTTTCAGGTGACGTACTTTGAGTTTCAAGATATCTAGACGCACTCTTTC 1440  
 Db 31172 ATATGCGGTTTTCAGGTGACGTACTTTGAGTTTCAAGATATCTAGACGCACTCTTTC 31231  
 Qy 1441 GTGAGACGCAATTCGCAAGACCGTCCGGAATAGTATACGTTTCTTTAAACACGCTACTG 1500  
 Db 31232 ATCGAAAGTCAATTCGCAAGTAAACCCCGCTTATGTGTACAGTTTCTTAGTACCTGCTA 31291  
 Qy 1501 TACAAAGCCATGTAATCTGTTTACGCGTACGCGGTGACTAGGTCAACAGAGAAATATCTG 1560  
 Db 31292 TATAATGCCATGTAATCTATTTAAGCGCACGCGGTGAAGCAGGCGCATGAAGAAATTCCTC 31351  
 Qy 1561 CAGACCTTAAATTTGACACCTCCCGCGCTCTGCAACCGGCGGTGTGAGACCTTCAA 1620  
 Db 31352 AGGACCTCAGGAACTGCGCGGTGTCTAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 31411  
 Qy 1621 ACGGTTTCGTGAAGATTAAACCTCGAAGACGACATCTTCTTACGAGTGTGCACTCCCGG 1680  
 Db 31412 GTTCTGGGCGGAGGAG-----GGGGAACACAGATGCTTCTTACCTCAACCGCGCATCCCA 31465  
 Qy 1681 CCACCGCGGATATCACTCCCTCCAGGTTTTGTGTCAGACGCTACTGCGCTTAAAGGAC 1740  
 Db 31466 CCATCCCCACCCACGCGGTCTTCCAACTGTGTGCGCTTATTAATCTCCGGGGCGAAGAA 31525  
 Qy 1741 GTGTTGCGCTCCAGAAATAAGGTGGCGTGTGCTGTGTTTGGCGGGGAGTTTGAGAAGAA 1800  
 Db 31526 CTATTTGCAACAGGCTGGCGCTGGCACGCGGACACTTTGTGACGAGTTTCTCCACTCC 31585  
 Qy 1801 ACGTTTCCGGTTTACGGTTAAACATGTTGTAAGGACGAGTTTGAATTTGTTCTCCCT 1860  
 Db 31586 GATTTTTACAGTTTACGGTGAACATCTGTGTGCGAGATGCGGTGGAATTTGTGTCACT 31645  
 Qy 1861 TCAGAACCTCTCAACGGGCTGTGGCGTTTGCATCGACCGTTGAATCGTATATAAATTAAG 1920

Db 31646 TCCCGCGGCTCCACGGTCTAGTGGCATACGATCCACTATAGACACTATATATATCCAG 31705  
 Qy 1921 GGTGTACAGTTTACCGGTAGCGTTCGGTGTGTCAGGGTTTTCAGGCTTTCAGTGTATGAC 1980  
 Db 31706 GATATATAGTTTCTCCAGTGTAGATTTCGGCGTCCAGGAGGACACGCTTCCAGGAGGAC 31765  
 Qy 1981 CTCAGGAAGAAGATGCGCTTCCCTGCTGTGTGCGAGACTCTAGCGGTTTATTCGCTGCTTA 2040  
 Db 31766 CTGGCGAAGAAAGATGCGCTTCCATATAGTTTGTCCAGGACTCATCGGGTTCATTTGCTGCTG 31825  
 Qy 2041 GAGATATACATAACCAATTCACCAACCATGAGGACGCGGAGGTTTTCACAGTGTGC 2100  
 Db 31826 GAAATATACGTCCACAGATGACAGAGACCTCGAAGTGGCGAGCTGTTTAAACATATGT 31885  
 Qy 2101 TGTGCGGGGACTATGGGCTCAGTCAAAATTTAGCCATGACCATCTGTAAGAGGACACAGGA 2160  
 Db 31886 TGTGAGGGGACTACGCTATCAGTTCTTAATCTGGCTATGACCATAGTGAAGGACACAGGG 31945  
 Qy 2161 ATGCTGTTGGAGCGAGTAGCCGTAGTATTTGGGTCCCAAGAAAGCTCCACAGAGCCAC 2220  
 Db 31946 GTTTCATTAAGTAGGTGGCCATATCGTTTCGGCAACCCGCAATATCAGAGCCAGTCTA 32005  
 Qy 2221 GTGTATGTAGCAATATCAAGGCTGTTAACTCAAAATTTTGGTCATGGAACAGCAACCCC 2280  
 Db 32006 GTGTATGTGGGTGTATCCAGGGCCATCGACGCTGTTACCTGTGTAATGGAAGTAATCCC 32065  
 Qy 2281 CTTAAAAACCTTCTCAGAAACACGTCGATTAACCTCCGCGCAAGCATATAGTCCGCGCC 2340  
 Db 32066 CTTAAGCTAAATGGACCGCGGTGACGCCGAGTCCCACTCTCAAAAGTACATCAAGGCC 32125  
 Qy 2341 CTCCACAAACCCAAACACAAACCTCATCTACTAA----- 2373  
 Db 32126 CTATGCAACCCAGAGACTTACTCTGATCTACTGACCGGTACCCCTCTCTTAGGACACTGAT 32185  
 Qy 2374 -----ATAATTAATAAACAAGCAANTGGAACAATTTGTG 2406  
 Db 32186 GTGTTTGGGAATAAAGCATGAGACTTGACACTTAATGTGCTGTATTTGACACCATTTCT 32245  
 Qy 2407 TTTTATTTTCACTTCCAAACACCGCCAGGATAGTTGTCAATTTTCCACACACCGGGGGGTGG 2466  
 Db 32246 TTAATTTTCACTTCCAGCCAGCCGAGTTATGACCGTTTCCACACAGGGGTGCGGTGG 32305  
 Qy 2467 CGCGCATATGTTGACGACAGAAACGCGCGGCGCTCTGGGTGTAAGTATGCCCC----- 2521  
 Db 32306 AGGCCAGGATGCGGGTTGGTGTGACCTGTGACCTGACCGCGGTAGTTGTGTTCTGTGATG 32365  
 Qy 2522 -----GCCCCCGGTGTGAGGTCTACTCTTTAGGCGCTGCGGT 2564  
 Db 32366 AAATCGAGTGGCGGAAGTACTTGGAGATTGGTTGGAGGTGACCTTTGTCTCGAGC 32425  
 Qy 2565 GTACCAACATATAATCAGCTCGTCACTCCGTTGCGAGTCCCTGCGTGAACGCGGCTCTTG 2624  
 Db 32426 GAGACACGATCAGCTCAGCGCGGACGAGGGCTCTCTGTGTGTCTCTCCCGGAGATA 32485  
 Qy 2625 CAGGCGGACGAAACACCGCGCTTGTAACTGATCT-----GCAT 2664  
 Db 32486 TAAATATCAGGAGCCACTGCTTTGCGGCTTAAAGTTTGGTTGTCTCTGCGAGGACCA 32545  
 Qy 2665 CGTCTGCTGTCTGAAGATGAACACAGGTCAATCACTAGTGGCGCGCGTGTGCTCAT 2724  
 Db 32546 CATCTCGTACACAGAGGCGGTGAGACTGCTTTTGGCTTCTGCGCCACGCTCCATGA 32605  
 Qy 2725 CGGCGTGTATTCGTCCTCATGTCACAGAGCTTCCGTGCTGCTGCTGCTGCTCATCTG 2784  
 Db 32606 GCCCGAATTTCTGACTCAATATCTCCCTTGGTCTTCTCCGCTCTCTCGGACGAGGGTG 32665  
 Qy 2785 AAATAATAATGCTCTTGTCTGTGCTGAGTTGATCCAGCAACACAGGTTTCTTCTCATCGG 2844  
 Db 32666 GCTGTGTGGGAAAAATGGCGCGCTCGGTAAACGGGCGCTCATTTGTTACGCTCCGAGAGT 32725  
 Qy 2845 AAGACATCCCGCTCTCTGCGGTGATCTAGATATATCACTGACAGGCGGCGGAGGCC 2904

Db 32726 TGGAACTGTCATCGTATCAGAGTCCGATGTCAGGTCGACGATCGCGTGGTGCAGCGC 32785  
Qy 2905 G-----2905  
Db 32786 GCAGGGGGCCACAGAGGCCCTTCATCAGGTCGCTGTATGTGTAACCTTTGTGTTCCAG 32845  
Qy 2906 -----TGGGAAAGCGCTCAGCTTTCGCGCACCTTCATCACTAGTGT 2946  
Db 32846 GTACACTATTTCGGAAGCAGGTGAAAGTCCGTATGCCCCGGTCCCACTGTATGCGGCCA 32905  
Qy 2947 CCTCGTAGTCACTATAAATAATCGCTTCGCTGTCTGATTTCAGATGACTGGGTGCGC 3006  
Db 32906 TCGGTTCCAGATAGCAACCCCTCGCTGCTGAAAGGTGAGAGCCCGAGCGGGAATC 32965  
Qy 3007 GGTGTTGGCCAGTACCAACCGCAACGGGC-----CGCGACCAACCAACGGC 3053  
Db 32966 CGTCATCCTGACTAACCCATCCCATGACGCTCGGACTCGCGTGTGTAAGTGC 33025  
Qy 3054 GTCACTGGCGGGGGCGCTCGGTGTGTGGTGGCACCTGTGAAATGACGGTAGGT 3113  
Db 33026 GCACGGCGCGCTACCACTGCTACCGGTTTGGGCGTATGGGCGGCTCTGGCCAGAGGC 33085  
Qy 3114 AGCGGGCGCTCTGTTAAAG-----ACTCACCGTCCCGTG 3150  
Db 33086 TCGGGCGAGTGAGATAGGTTGAAAAGTCTGCAGGAGTACCCCTCTGGCTCGCTT 33145  
Qy 3151 CGTCGGATCTATTCTAGTGGGAGGAGACCGCGGTTTCGTCTTCTTATCCACATCAC 3210  
Db 33146 CCTCCTGAACATCGTCACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 33205  
Qy 3211 TGTAAGTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3270  
Db 33206 ATTCGCGCTCGACTGATCCGGGATATCTGATATCAGAGGGTTGCTGGCGCGATG 33265  
Qy 3271 GGGAACTTTGTTCAAAACATCTTCAGTGCAGTCAATGATCACTGGGGGCTCCCGT 3330  
Db 33266 GCGGTCTCTCGGCAAGCGTCTGCTGGGAGACATATCTATCACGCTGGTCCAGCAT 33325  
Qy 3331 ACCGGGGGTACCGGCAACCCCTAGCGA-----CCGGCGCAACAAAC 3378  
Db 33326 AGCGCGCGGCTGCCAATCTCGAAGTATGAAAGAGTGGAGTGGGAATATGAAT 33385  
Qy 3379 GTCTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3438  
Db 33386 TCACGGGGGTCGCTCGAGGCGCTCTTCTCAATGGAAGCATCTCTTCTTCTTCTTCT 33445  
Qy 3439 GGGTCTTACGGGTCGGATGAAGAC-----ATAGCCATGGCC 3476  
Db 33446 TGCTAGACGAGTCTCAAAACATCGCCATGGCTTGTACGGGGTTGACCGCTAGGGC 33505  
Qy 3477 GAAAGTTCACCTGTAAACCGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGG 3536  
Db 33506 GGAATTTTAAAGCAGAGTTATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 33565  
Qy 3537 CACCCCTCCGATCTGAACTAGATGATTCGAACTTTAAATGTTTGCAGCCCTAGAA 3596  
Db 33566 TCTCAGCGCGTGGCGAGTCAATAGTCTGTTGGCTAGTTAAAGTATTACAGCCCTGGA 33625  
Qy 3597 ACTCGGCGAGGAGAACCGGTTCCCGTGGCGGTAGCAGGGGAGACGGATGGCG 3656  
Db 33626 ACCAGGCGCATCGGAGGTTCGGGCAACAGAGAG-----GCACGGGAGATGGATGTGG 33682  
Qy 3657 CTTGAGGAGCAGATGCTCTGTCGCTTTATCAGCGTCGCTTGTCTATAGCCCTTTCTTC 3716  
Db 33683 CCGTAAGCACCAGGTGTTCTGTCGTTTATGAGCGGAGTCTGTCAATGSCCTTGGCC 33742  
Qy 3717 CCGCAGCATAAACCGCACTCTGAGCTTGGCAGAAAGGCGCTTATTATGTAGTGG 3776  
Db 33743 CCGCAGGAGAAAAACGCAATGTTCTAACTTTGAGGATATGCTACTGATGATGAACCTCG 33802  
Qy 3777 TGAACCAATCCAGCAGGTTGGCTGTGACCCCGCTTCCCTTCTTCCACCGTCAAGAA 3836  
Db 33803 TGAACCAATCCAGCAGGTTGCTGCTGTGAGCGGCGCTTCCCTTCTTCCACCGTCAAAA 33862

Qy 3837 TGGTGTTTAAAGCAGAACTCCCGTTCCTTTCGCAACAGTCTTAAGCACCCGTCGAGAAGGAG 3896  
Db 33863 CTGTGTTTAGTAGCACACACCCCTGGCAGCCACCTGTCGAGGACACCCGTCGGAAGGAG 33922  
Qy 3897 CGGTGAACCCAGGTACAGTATTGGCTATCTCTTTTAAATAATTTTGGGCTTGGAGGTA 3956  
Db 33923 TACTGAAATTTGGGAGCGAAGCCTCTAGCTCTCTTAAAGATGCTTCTCAAACTGGGTGAA 33982  
Qy 3957 TACTGTAGTCGGAGCAACCGCTAAAGGTAGTCCGGTGGCTTGACCGGATGGTAAGGAT 4016  
Db 33983 CCTGACATTCGGATCCACACTTAACGCGCAGGCACTAGTTCGCCCTTGTGTACGGT 34042  
Qy 4017 CTTGCCCAAGGATCAGACTTTTAAATATCTCTCGGTTCGCAACAATACACACCACTAA 4076  
Db 34043 CTTGCCCTAAGATCACCACTTTTAAATATCTCTGATCGCAGCAGTGGGACCACTCA 34102  
Qy 4077 TTCTATCTATCGGTGATTAATCACGTTACGTTTCTCATATCCATAACGCGCTTCAGCA 4136  
Db 34103 GCTTGTCTGTGGGATACACTGTGGTGTAGCTTAAGTTCGGAATCTGTCTGAGCA 34162  
Qy 4137 GAGCGCGGAGCTTCTGTTTAAATAATGGAGAAAGATTTTAAATAATTCAGGCCA 4188  
Db 34163 GCGAGAGCAGTTCTGTTTTCAGAAATGATGAGAGGCTCAGAAAGGAAATCCA 34214

## RESULT 9

BD190715 35100 bp DNA linear PAT 17-JUL-2003  
LOCUS Unique associated Kaposi's sarcoma virus sequences and uses thereof.  
DEFINITION BD190715  
ACCESSION BD190715  
VERSION BD190715.1 GI:33000454  
KEYWORDS JP 2002513274-A/15.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 35100)  
AUTHORS Chang, Y., Bohenzky, R.A., Russo, J.J., Edelman, I.S. and Moor, P.S.  
TITLE Unique associated Kaposi's sarcoma virus sequences and uses thereof  
JOURNAL Patent: JP 2002513274-A 15 08-MAY-2002;  
THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK  
COMMENT PN JP 2002513274-A/15  
PD 08-MAY-2002  
PF 22-JUL-1997 JP 1998509105  
PR 25-JUL-1996 US 08/686349,25-JUL-1996 US 08/687253 PR  
25-JUL-1996 US 08/686350,25-JUL-1996 US 08/688814 PR  
25-JUL-1996 US 08/686243,05-SEP-1996 US 08/708678 PR  
10-OCT-1996 US 08/728323,13-NOV-1996 US 08/748640 PR  
13-NOV-1996 US 08/747887,29-NOV-1996 US 08/757669 PI YUAN  
CHANG, ROY A BOHENZKY, JAMES J RUSSO, ISIDORE S EDELMAN, PI PATRICK S MOORE  
PC C07H21/04,C12Q1/68,C12P19/34,C12N15/10  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.

## FEATURES

source  
1. 35100  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 11.0%; Score 1002; DB 6; Length 35100;  
Best Local Similarity 54.5%; Pred. No. 2,1e-253;  
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;  
Qy 1 ATGGAGAGTTCGTCGATGGACCAACACGTCGAAACCAATCCGGGTCATCTTGAAC 60  
Db 29792 ATGGACAGTTCGGAAGGTCGACTGATATGGACGAGCCCGGGTTCATCTCTAAC 29851  
Qy 61 ATGACGTCGATGCCAAAGTCAGGGGTGTCGTGGATCATGTCAGTCGCTGTCAAATA 120

Db 29852 ATGACATCGGATGCAAGTCAGAGCGTAGTGGACAGATTGACAGGTTGTCAATATC 29911  
Qy 121 ACTACAGCCACCGGAAATGGGTGGTACGACTGGCCCTTCGATCGGCTGAAGACTCC 180  
Db 29912 ACTACCTCGCCCGGAGATGGGTGGTATGACCTAGAGTTGATCCATCGGAGACGAA 29971  
Qy 181 GGGCCGTTCTGGCGTTTACCGTTTATCTAATACGGGAACTGCTGGTGTGGGAAAGT 240  
Db 29972 GGGCCCTTCTGGCGTTTCTGGCATAGTAAATACGGGAACTGCGAGAGCGGGGAAAGC 30031  
Qy 241 ACCAGCATATGGGCCCTGTACAAAATTTAAACTGCTGATACAGGGCGGACCAACATA 300  
Db 30032 ACCAGCGTATCGGCCCTACATCAGATCTCAACTGCTCAATATTACGGGGCTACAGTGGTA 30092  
Qy 301 GCCCAGACAGCTATCGCGCTGCTAAAGAGGTTCTGCCAGATCTTTCAGCGTTT 360  
Db 30092 CGCGCAGAGATCTTTTCAGGGCTTTAAAGTCTCTGCTGCTCCATATACAGCGCTTC 30151  
Qy 361 GGCTTTAAGAGCCGACATCAATATATAGCGGTGAGAAAGCTCATCAGACCGGAGCGGTA 420  
Db 30152 GGATTCAGAGCAGACATTAATCTGTCAGAGGAAAGTGGCCAAAGTAACTCAGTCC 30211  
Qy 421 TCCATAGAGCAATTCAGCACAGGAGCTATCGAAGTATTGGCCGGTTATAGTGGACAT 480  
Db 30212 TCCATCGAGCAACTCCAGAGATCGAGCTGGCTAGGTACTGGCCAACTGTCACCGATAT 30271  
Qy 481 ATGAAAGAGGTTATGGCGAAATAACCCAAATGGCATGTACGGGACTATATCCACGCGAAT 540  
Db 30272 ATTCGAGAAATTTATGGCAAGAAACAAAGGGCGAGTATAGTCCCTCTCTCAAAGCGCT 30331  
Qy 541 TTGAAACCTCTCGAGATGACCGGACCGGTTTATGGACTTCCATATATTGTATATC 600  
Db 30332 TTCAGACTCTTTCGCGTATGGGTGGAGCCAAATTTGGGACGAGTAACTATCGTGATA 30391  
Qy 601 GACGAGGCGGAACCGCTCTCTTACATATCTACACCGCTCGGTGTTCTTTTACTGGTTC 660  
Db 30392 GACGAGCTGAAACCTCTCTGCTCCATATTTTGACGGCGGTGGTCTTCTTATGGTTT 30451  
Qy 661 CTAACAGCTGGCTAAATACCCCTCTTTACCGCAGGGGGGGTTCGGTGATAGTATGC 720  
Db 30452 TACAACAGTTGGCTGGACACCCGCTATACAGAAATGGTGGCGTGCCTTGATAGTCTGC 30511  
Qy 721 GTCGGTTCGCAACGACAGCAAAACGCTTCCAGTCAACTTACAACACGAGGAGCGCAAAAG 780  
Db 30512 GTGGGCTCTCCACCCAGACGACGCTTTCACTGGCTTCAACACACGAGCAGAGA 30571  
Qy 781 ACGGAGATATCGTCTGGAGAACATCTTAATCATATGATAGGGAAGAGGTGCTATCT 840  
Db 30572 AACGAGATATCTGCTGTGATAATGTGCTCACTTCTTATGGGAAACGCTGAGGTTGCA 30631  
Qy 841 GAGTACGTACATTTGGAGAGAACTGGCGCTGTTTATAAACAACAGCGCTGCACCGAT 900  
Db 30632 GATTATATTAGGCTGGAGAGAAATGGGCCCTATTTATTAACAATAGCGCTGTACCGAT 30691  
Qy 901 CTGCAAGTTCGACACCTTTCTAAAATTTTAGAGTATAATCTTCCCATCTCTGACGAAAGTC 960  
Db 30692 CCCCAGTTGTGCTACTTCTGAAGACCTTAGAATATAATCTAGACATATCACAGAGTTA 30751  
Qy 961 ATGAGTTACGTAGACAGATTTGTCTCTTCTTAAAGTATAGATATGGATCTTTTAAATAC 1020  
Db 30752 ATGGACTATATAGATAGTTGTGGTTCGAGAGTAAAGATCTTGGACCCGCTCGATAT 30811  
Qy 1021 ATTGGCTGACCCGACTCTTTTGTACATAGCGAGGTAAGGCGTATCTTAACAAACCTA 1080  
Db 30812 GCAGGTGGACAAGACTCTTCTATCTCACACGAGGAGTGAAGTCTTTCTGGCAACGCTG 30871  
Qy 1081 CACCATGTCTAACGCTAGGGGGGATACCGAGGACAGAGGCTTTTACCTGTCCCGTG 1140  
Db 30872 CACACTGCTCTCGAGTAAATAGGATGCTGTGTCCAAAGCTTTTCCCTGCCCAGTG 30931  
Qy 1141 GTGTGCGAGGTGTTGTGAAGCCGTTTGAAGAAATACAGCGGGCGCTCAACCTCACCAAC 1200

30932 GTCTGTAGGTGTTTTACAGAGCCATTTTGAGGAGTACAAACGGCGGTAGGCTCACACAC 30991  
1201 CTCACCGTGACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAATTTTCGAGTTT 1260  
30992 ATGACTCCCATAGAAATGGGTAAACAAAAATCTTTTCAGCTAAGTAATCTCGAGTTT 31051  
1261 GTGGACCAAGCATGTTCATAGTGGCCACGGAATCTCACTGAACGTTTCGACCCAGGTTACC 1320  
31052 GCTGATCAGGACATGGCTGTGTGGGACCTATATACAGACGGGTCCACACAGATACC 31111  
1321 TTATATCCAAAGTTGTTTAAAAACAGCCACGCTGTCTTTAAATGGAAAAACAAAAATGC 1380  
31112 TTGGCCACTAAATTTGTCAAAAACAGCTATGCTACCTTACTGGAAGACCAAAAAATGT 31171  
1381 ATATGTGGGTTTTCAGGGTACGTACTTTGAGTTTCAAAAGAAATCTCTAGACAGGAACTCTTC 1440  
31172 ATATCGGGTTTTCAGGGTACATACCAAAAGATCAAGTCCATCTAGACGGGAGCTATTT 31231  
1441 GTGGAGACGCAATTCGCAAGACCGTCCGGAATAGTATACGGTTTCTTAAACACGCTACTG 1500  
31232 ATCGAAAGTCAATTCGACAGATAACCCCGCTTATGTGTACAGTTTCTTAGTACCTGTGTA 31291  
1501 TACAACGCCATGACTCTCGTTTTCACGCTACCGCGTACGTAGTGTACACAGAGAAATATCTG 1560  
31292 TATAATGCCATGACTCTCATTTTACGCGCACGGGGTGAAGCAGGGGCATGAAGAAATCTCTC 31351  
1561 CAAAGACTTAAATTTGCAACCCCTCCGGCGCTCTGGCAACGGGCGGTGTAGACCTTCAA 1620  
31352 ACGGACCTCAGGAAACTCGCGGTGCTCAAGAGCTGATCTCTGATGAGCTCCAGGAC 31411  
1621 ACGGTTCTGTAAGAGTTAAACCTGGAAGACGACATCTTTACACCGTGTGAGTCCCGG 1680  
31412 GTTCTGGGCGAGAG-----GGGACACAGATGCTTCTACTCACCGCAGACCTCCA 31465  
1681 CCACCGCGGGTATCACTCCCTCAGGTTTGTGTCGACAGCTACTGCGCCCTTAAGGAC 1740  
31466 CCATCCCGCCACCGCGGCTCTTCCAAACACTGGTGGGCTATTACTCGGGGCGCAAGGAA 31525  
1741 GTGTTGCGCTCCAGAAATAAAGGTGGCTGTCTGCTGTTTGGCGGGAGTTTGAGAAGGAA 1800  
31526 CTATTTCTGCAACAGGCTGGCCCTGCACGCGCACATTTTGTGACGAGTTCTCTCACTCC 31585  
1801 AGTTTTCCGCTTTACGTTTAAACATGTGTCGTAAGGACGAGTTGACTTGTCTCTCCCT 1860  
31586 GATTTTCAACGTTTACGTTGAACATCTGGTGGCAGATGGCGTGGACTTTGTGCTCAT 31645  
1861 TCAGAACGCTCAACGGGCTGTTGGCTTTGTCATCGACCGTTTGAATCTGATATAAAATTAAG 1920  
31646 TCCCGCGGCTCCACGCTCTAGTGGCATACGATCCACTATAGACACCTTATATATCCAG 31705  
1921 GGGTACACGTTTTTACCGGTACGTTTGGTCTGCTGTGAGGTTTGGCTCTGATGATGAC 1980  
31706 GGATATACGTTCTCTCCAGTGAGATTGGCCGCTCCAGGAGGACAGCGCTCAGCGAGGAC 31765  
1981 CTGAGAAAGATGCCCTCCCTGCTGTCGAGGACTCTAGCGGTTTTATCGGTGCTCTA 2040  
31766 CTGGCAGAAAGATGCCCTCCCATAGTTTTCAGAGACTCATCGGGTTTCAATTGCTGCTG 31825  
2041 GAGAATAACATAACCAAAATTGACCGAAACCATGGAAGGACGGAGCGTTTTTCAAGTGTGC 2100  
31826 GAAAATAACGTCACCAAGATGACAGAGACCTCGAAGGTGGGACGCTGTTTAAACATATGT 31885  
2101 TGTGCGGGGACTTATGGGTCAGCTCAAAATTTAGCCATGACCATCTGTAAGGCACAGGGA 2160  
31886 TGTGCGGGGACTTACGGTATCAGTTCTAATCTGGCTATGACCATAGTAGGACACAGGGG 31945  
2161 ATGTCTGTGGAGCGAGTAGCCGTAGTATTTGGGTCCCAAGAAACGTCGACAGACCCAC 2220  
31946 GTTTCATAGTAGGTGGCCATATGTTTGGCGCAACACCGCAATATCAGAGCCAGTCTA 32005  
2221 GTGTATGTAGCAATATCAAGGGCTGTAACTCAAATTTTGTGTATGTGATGACACACCC 2280  
32006 GTGTATGTGGGTGTATCCAGGGCCCATCGACGCTCGTTTACTGTATATGACAGTAATCCC 32065



RESULT 10	EHVU20824	184427 bp	DNA	linear	VRL 02-FEB-1996
LOCUS	Equine herpesvirus 2, complete genome.				
DEFINITION	U20824				
ACCESSION	U20824				
VERSION	U20824.1				
KEYWORDS	GI:695172				
SOURCE	Equine herpesvirus 2				
ORGANISM	Equine herpesvirus 2				
REFERENCE	Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae.				
AUTHORS	1 (bases 1 to 184427)				
TITLE	Telford, E.A., Watson, M.S., Aird, H.C., Perry, J. and Davison, A.J.				
JOURNAL	The DNA sequence of equine herpesvirus 2				
MEDLINE	J. Mol. Biol. 249 (3), 520-528 (1995)				
PUBMED	95302501				
REFERENCE	2 (bases 1 to 184427)				
AUTHORS	Telford, E.A.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-FEB-1995) Elizabeth A.R. Telford, Institute of Virology, MRC Virology Unit, University of Glasgow, Church Street, Glasgow G1 5JR, United Kingdom				
FEATURES	Location/Qualifiers				
Source	1. 184427				
	/organism="Equine herpesvirus 2"				
	/mol_type="Genomic DNA"				
	/strain="86/67"				
	/db_xref="taxon:133899"				
repeat_region	1. 17553				
	/rpt_type=direct				
	9248..10399				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..10499				
	/note="ORF E1"				
	/codon_start=1				
	/product="G protein-coupled receptor"				
	/protein_id="AAC13788.1"				
	/db_xref="GI:695173"				
	/translation="MATTSATSTNTSSLAATTTMTNTFTSLTSSVTTIASLVPSTNSS				
	EDYDDLDVDYSESAPCYKSDTTLAAQVVPALVLLVFLGLGILNVLIVIRYMK				
	IKNTLNLLAIISDLFLTLTPFWHYIGMDWTFGISLCKLRGVCYWSLSQV				
	FCILLVDVLAIVATLRTVTTCGIVTCVTFGLAGLISLPEFFHGHODDG				
	RVQDDPYPMSTNVRRAVAKVIMLSILPLIMAVCYVILIRLLRPSKKYKA				
	IRLIFVIMVAFVFWPTYNIVLLSTFPHATLNLQALSNLMDMLLIKTVAIYHCC				
	INFIYAFVGEKFERHLYHFHTYVAILYLCYIFPLSGDGEKGPTR1"				
	10494..1				



LCSSLLPCLRLQNDTSCSPLSLELOHLLQNTREALGVLTDLLSGDPSRSEYFEALHPS  
RPLEGCKRHARVFPFYGAETKSVLSLNDVEFFFKELNSVFCVCLPAEGALEALGE  
TVAFGLKRGVSPFPADAYSSVPCASGCFAPAAMLPNQGESVLSMLAIVNCVNCRCQ  
VPDPVIGVFNFEHLHGLADARAAGRGRESERRREDADDEDEEDDEEDDEEDGGDA  
GDAALVLTTESSLVAGHTIPEEDGRILAIISNLVYSSAAADRGVGGVGTARATSSS  
HMAKLAHARWERSPAWLGRAQSPFPAHPSLFCG3VFNISDDTAAALQKD  
CSATFLKSNQYPLIOQNELVYRLNEVLNGRDEGAKGAEATADEAPLKPQDGA  
CDPLSDARVARDLYLKKLTKDGLRLRLDCIETHGRVLSTLSLRVWGSALYASAA  
RLVNHLEFRFQVGLGADLTAGGEAAAFENSKYIKNALHGQNLREHLDLSIVVHYRL  
ITGPLSLQNSHPFPVNPVALAYCLDAAGAMPHQKLVITEMIPWGESKDWIDCNFNSP  
YSTETGLMLTKTNTYIREAVLSISLYNRWERSKLSLSATLRSGLSCLABSAGSEL  
GEGVLYTGETAPLVLFVDSKGVFKDLVTLVLYTHLQLSGRQASV"  
polya\_signal  
30114..30119  
/note="ORF E3, 03, E4 and 06"  
/notes="ORF 03"  
/codon start=1  
/product="glycoprotein B"  
/protein\_id="AAC13795.1"  
/db\_xref="GI:695180"  
/translation="MSVGGPRVVLCLVCAALLCGVAQEVVAETTTTFFATHRPEV  
AENPAPLFFRVFGASPTGGEIFFPLESCPNTEKDHIEGLAIYKTNIVP  
NVRKYKMTSTIYKGBEDATINQHTSYAVPLYEQMDHYOCFSAVVNEGGH  
VNTYDRDNENAFILKPADGLTSSITRYQSOPVATPRNLLWSYTRTTVNCVTE  
MSARMKPEFFVTSVGDITENSPLFKNGTEPEKILKRPHSIQLLKNVAVTKYVGL  
GQADNATRPALFGDYSLSWKAATTENSSYCDLILWKFPSNAIQTOHNSLHPANDIT  
ASSTPLERENEPEKCIWNNTOREIKKLEVEKTRPNGTAKVYKVTGKNLYVM  
QPLQIDLDTHAKLINLNAVASPTSTETTPRRRRRTSSVSGGNGDSTKES  
VASQVQFVNRKINYLNLGELSRWCREQTRASLMWYELSKINPTSVMGAIYRP  
VSAKLGVVSDCTSVQSVFVKMKVPKEDLCYTRFVGFKINGSELFAQ  
LSPREVIDTSVQSCHEHYFOAGNQMYKDYVYVSTLNTLDIPLTHMITLN  
LSVENIDKAVILEYKTEKRLSNVFDIETPREYNYTONLNGRLKLDSDIDHGRD  
SFTQTLGDMODLGTGKVVANVASGVSLFGISVSGVLSFFKPNPFGMLLVLIAG  
VVVYLPMTRESIYSAPIEMLYPGVERAAQEPGAHPVSEDIIRNLHMHQFOQRQ  
AEBEARREVEKGRILFEVIRDSATSVLRRRGGGYORLQDGDGDDDEGDEYPLRRQ  
DGGYDDVDVEAGTADTVG"  
polya\_signal  
34855..34860  
/note="ORF 07 and 08"  
/notes="ORF 07"  
/codon start=1  
/product="DNA polymerase replicative subunit"  
/protein\_id="AAC13796.1"  
/db\_xref="GI:695181"  
/translation="MSGFNPYLVRKTLFLKKAAPSRPTKEYTRIIPKCFKTPGAAGVVP  
HTSLDPCVPGVGGKTEPIYLDGSRSLWAGRGPGTGAGGQHTPVALTFHYDIE  
LDTVLQCCRENEFPQFOTDIIIPSGTVLKLGRSTDORSVCNVFROELFYVVRPSELK  
LDPLIQCCRENEFPQGRYVEKTSKVLREVCVAREVYVVRASSGQFVLLDAGEL  
TAAGCEVFEINVDARFIIDNGFTFGYSCAAVPROGGAARSWTELEVDCAAD  
LEPHAGRMFGNVLVSFDIECENGFPNARDMDILQISCVLWAGSGEAPRSVL  
LNIGTCEELEGVEVYQPSBELDLILFFTMIRDADEVFTGYNINFDFFYVIDRATQ  
VYNLKEFTRVASSSIFEVHKPNSAGFMRAVSKVAVGVVPMIDMYQVCRDKLSLS  
NYKLDTVAGCEKAKEDVSYKEIPLHFRQPGGRARILGLYCVKQSAVLVLDLRYFMT  
HVRISLAKIAKTPTRVLTGQOIRVFSCLLDVAGREGVILPVRHADAEYQCATV  
IDPSRFGVNTPLVDPASLYPTLIQAHLNLCYSTWIPGDELCLHPLHGPQDYETELA  
SGPVHFKRHKVSLIATLLNWLAKRKAIIRLRELTASDEAVRTILDKQQLAKVTCN

Query Match 10.7%; Score 974.6; DB 14; Length 184427;  
Best Local Similarity 54.2%; Pred. No. 5.2e-246;  
Matches 2289; Conservative 0; Mismatches 1859; Indels 79; Gaps 12;  
Qy 34 GAACCAATCGGGGTTTCATCTTGACATGACCTCGATGCCAAGTCAGGGGTGCTCG 93  
Db 92963 GAGCCGGGGCGGGGTTTCATCTTGACATGACCTCGATGCCAAGTCAGGGGTGCTCG 93022  
Qy 94 GATCACCTCAGTCGCTGTCAATATATACTACACGCCACCGGAAATGGGTGGTACGAC 153  
Db 93023 GGCAGGATCGGGGCTGGCGGCAAGACCGTCCCCCGCCGAGATGAGTGGTTCGAC 93082  
Qy 154 CTGCGCTTCGATCGGCTGAAGACTCGCGGCGCTTCTTGCCTTTACCGTTTATCTAAT 213  
Db 93083 TCTCAGTTCGACCTCGAGGCGGGGCGGCTTCTTCCCTTCTCCACCTTCTCTGATA 93142

Qy 214 ACGGAACTGCTGTGTGGGAAAAAGTACACAGCATATCGGCGCTGTACCAAAATTTAAAC 273  
Db 93143 ACGGCAACGGGGGGGGGGTAAAGAGCACTAGTATCTCGGGCGCTGTATCAGAGCCTCAAC 93202  
Qy 274 TGCCTGTATCAGCGGCGGACCAACCATAGCCACACAGAACCTATCGCGTCGCTTAAAGACG 333  
Db 93203 TGCCTCATCAACGGGGCCACGGCCGTGGCGCTCAGAACCTGTCCAAAGGCTCAAGACC 93262  
Qy 334 TTCTGTCCACGATCTTCAGCGCTTTTGGCTTTTAAAGACGACACATCAATATATAGCGTC 393  
Db 93263 TACTGCCCCACCATCTACAGCGCTTCGGGTTTAAAGACGAGCACTTAATATATATGCGC 93322  
Qy 394 AGAAAGCTCATCAGACCGGAGCGTATCCATAGAGCAAAATTCAGCAACGAGAGCTATCG 453  
Db 93323 CGCCACGGCCGAACCGCGCCGCGGACATGAGGGGATCCAAAGGAATGAGCTGTGC 93382  
Qy 454 AAGTATTGGCGGTTATAGTGACATTAATAAGAGGTTTATGGGAAAAAACCCCAATGGC 513  
Db 93383 AAATACTGGCCGTGATCTCAGATATATCTCGGGAGTTTACCAAAAAGAGCAGCGGGGC 93442  
Qy 514 ATGTACGGGACTATATCCAAACGCGAATTTTGAACCCCTCTCGAGNATGACCGGACCGTGT 573  
Db 93443 CAGTACGAGCACCTGACGGGGCGGCGCTTTGGGGCGCTGGCAGATGGGACACCCCAACC 93502  
Qy 574 TTATGGACTTCCAAATATATTATGTAATCGACGAGCCGGAACCCCTGTCTCTTACATATC 633  
Db 93503 CTGTGGACCAACCAATCATTTGTATAGACGAGGGGTACCTGTCTCGACATCTGT 93562  
Qy 634 ACCACGCTGTCTTTTACTGGTTCTTAAACAGCTGGCTTAATACCCCTCTTTACCGC 693  
Db 93563 ACGCGGTGTATCTCTGTACTGTGTTTAAACAGCTGGCTGCGAGACCCCTCTGTACAG 93622  
Qy 694 CAGGGGGGGTTCGCTGCATAGTAGTTCGTGGTTCGCCAACGACGAGCAAAACGCTTCCAG 753  
Db 93623 AGCGGGCGGTGCTCTGGGTGTCTGTGGGTGCGCCACGACGAGCGGACCATACAG 93682  
Qy 754 TCAACTTACACACCGGAGCAAAAGACGAGATATGTCGTGGCGAGAAATCTCTTAACA 813  
Db 93683 TCCACCTACAAACCACTCCATGCAGAAACACACATACAGAGAGTGCAGCAACATCTCTACC 93742  
Qy 814 TTCTATAGAGGAGAGGTCGTATCTGAGTACCTGACATTTGGAGAGAACTGGCGCTG 873  
Db 93743 TTTCTCATGAGACGAGGCGGTGTCCAGGTACCTGGACCTCAACCACTTGGCGCTG 93802  
Qy 874 TTTATAACAACAAGCGCTGCACCGATCTGCAGTTTCGAGACCTTCTAAAAATTTAGAG 933  
Db 93803 TTTATCAACAACAAGCGGTGCACGACCCAGAGTTCCGGCACCTGTCTCAAGACCTCGAG 93862  
Qy 934 TATATCTTCCATCTCTGACGAAAGTATGAGTTACGTAGACAGATTGTCTCTCTAAA 993  
Db 93863 TACAACTTGACATCTCCCCCGTATGATAGACTACATCGACCGCTTCTGTGTGCCAAG 93922  
Qy 994 AGTAAGATTATGGATCTCTTTAGAAATAATTGGCTGGACCCGACTCTTTTGTTCACATAGC 1053  
Db 93923 AGCAAGATCTCTCCCCCTCGAGTACGTGGGCTGGACCGGCTCTTCTGTCTCCACAGG 93982  
Qy 1054 GAGTAAAGCGGTATCTTAAACAACTTACACACATGTCTAACGCTAGGGGGGATACACAG 1113  
Db 93983 GAGTCAAGGCTTACCTGACGCGCTTGCACGAGACCTGTCCCTGAAACCAAGGTTGGCGC 94042  
Qy 1114 -----GACACGAAGCTCTTTTACCTGTCCCGTGGTGTGGAGGTGTTTGTGAAG 1161  
Db 94043 ACCCGGAGGCGGACGCGAGGCTCTTACCTGTCCCGTGGTGTGGAGGTGTTTACCGAT 94102  
Qy 1162 CGTTTGAAGAAACAAGCGGGCGCTCAACTCACCAACCTCACCGTACCGAGTGGGTG 1221  
Db 94103 ACCTTCAACGAGTACAGGAGCGGTGAACCTGCGGGGCTCACGGTCAACGAGTGGGTG 94162  
Qy 1222 ACMAAARACCTTTTAAAGTTAAATTTATTCGAGTTTGTGGACGAGCATGTCCATA 1281  
Db 94163 CAGAAGAACCCTTTCGAGGCTCAGCACTACTCCAGTTTATAGACGAGACTGTCCGCC 94222  
Qy 1282 GTCCGCCGGAATCCACTGAACGTTTCGACCCAGGTTTACCTTTATACCAAGTTTGTATAA 1341

95303 CCAGACACATACCTCTGTTTACTGAATGGAGATA-----CCACTACAAAAACCTTTACAG 95353

2410 TTATTTGAGTCCAAACGCGCCACGGATGTTGTCATTTTCCACACACCGGGGGGTGGCGG 2459

95357 ATTATCTGTTATTTGCTTGGCAATAAAAAACAATAAACTCAAAAAATCTGTTAAATTTTT 95416

2470 CCATAGTTTTGACACACAGAAACGCGCGGGCCCTCTGGGTTAGAAGTAGCGCGCCCGCC 2529

95417 TTTATTTTTTATTAATTAATCAATGGGAGAGAGATAGAGAAAGAAAAAGCCCTTAG 95476

2530 TGGTCGAGGTCAACGTCCTCTTAGGGCCCTGGGTTGTACCAACAATAAATCACTGTCAT 2589

95477 AGTTTTACGCCAAGCCAGTT--GTAGGGTGGGTTCCCGGCCACAGAGGGCGCTGTCGCG 95534

2590 CACGGTTGAGTCCCTCGGTGACGGCGCCCTCTTCAGGGCGGCACGAACACCGCGCGTTG 2649

95535 CA-----TATAGTGTTCGCATGACTCTCTGGGACATGACCACAGGACCTGATCTCTGGGCC 95590

2650 TAACGTGATCTGATCGTGTGTGCTCTCTGTAAGATGAACACAGGTCAATCAATCACTAGTGG 2709

95591 CAAGAGGCTCCCTCCCGGCGCGCACCCCTCTCCCTGCTCGGACTCAACGGAGG 95650

2710 CGCCCGTGCCTCATCGGGGTCTGATTCGTCGCATGTCTCAACAGAGCCTGGTGTCTCA 2769

95651 AGATATAGATCACTTTGGTGTGTTGGGTC-----ATTGTCCAGGTGCTGGAGGTGT 95704

2770 CTTCGTCGTCTATCTGAATAATAATAGTCTCTTGCTCGAGTGTGATCCAGCAAAACAG 2829

95705 GCTTTCTCTTTGAGAAAGCGGCGCGCAGCGCGCGCAAGGGAGAGGCTCGCCCTCGC 95764

2830 GTTCTTCTTCATCGGAAGACCATTCGCCCTCTGCGCGTGATCTAGATATATCACTGACC 2889

95765 TGTCTGACTCTCCGAGGA-----GAGCAGGCTGATGACAGGGACTTCCTCCTTTGGCC 95817

2890 AGCGCGCGGAGCGCGTGGCGAAAGCCGCTCAGCTTCCCGCACCTTTTCATCAGTGTCT 2949

95818 GGAGGGGGTTGGCGCGGGCGGGGAAAATTCCCGCGCCAGCGAGCTGTGACCTGCTC 95877

2950 CGTAGTCTATCTATAAAAAATCTGTTGCTGTCTGAATTCAGATGACTGGGTGCGCGGT 3009

95878 AGGATATCTGCTCGGAGTGTGCGATGCCATCGCTCTCGCTGTGCTGTGCTGTG 95937

3010 GTTGGCAGGTAAACGCGCAACGCGCGCGGACCAACCAAAACGCGCTCAGTGGCGGGCG 3069

95938 CTTTCGCTATCCGAAGAAGAGGAGGAGGAGTGTCCGAGATCTCTGACATG--T 95995

3070 CGCTCGGTGTGTGGTGGACCTGTCCGAAGTAAATGACGTAGGTAGCGCGCGCTCTGT 3129

95996 AGCGGGCATCTTTATCCCGCCGCTGCGCGTAAAGGGAAGTTTTTGGTCATGTCATCGC 96055

3130 TAAAGGACTCACCGTCCCGTGTGCGGATCTATTCTAGTGGGAGGAGACCGCGCGTT 3189

96056 TGTCTGCTAGAGCCCCCGGGGAGTCTCTGAGAGTGCAGCGTCTCTCATCGTCT 96115

3190 CGTCTTCTTCATCCACATCACTAGTTCGCTGTCTCCGAATGAGAGTCGGTTTGTCTG 3249

96116 CCCCCCGCTCTCGAGGACCTCGCTGTGCTGTATCGGACGCGCCCGGACCGCTCC 96175

3250 CGCTTTTGGAGTAGTGGAGGGAAATCTTGGTCAAAAACATCTCTCAGGTGACGTATGT 3309

96176 GCGGTTTTGGCTTACTAAAAACTGGAACCTGGGAGAGGGGGTGTGCACTGGGTGCAGTCT 96235

3310 CAATCACTGGGGGCTTCGTAACCGGGCGGTACCGGCAACCCCGTACGGGACCC----- 3364

96236 CAAAAAAGACATCATCAGGTCCCGCGTCTTTAGTCTTCTCATAGTTTCCCTTAATCC 96295

3365 -----GGGGAAACAAAAACGCTTGTGTTCTCTTTTCTTAGTGTGCCCGGAATCGGCA 3417

96296 CATCTTTGGGAACTGAAAGTAGGTGCCCGCTTTAGCTGGCGCCCCCGCGGTCTG 96355

3418 GCATCTGGGAGTGGCC---GGCGGGTCTTTCAGCGGTCCGATAGAACATAGCCATGG 3474

96356 TCATCTCTGGTGGCGACCGGTCGCGCCCGCGCTTCTCTCCGAGGACCCCTCTAG 96415





QY 481 ATGAAAGAGTTATGCGGAAACCAATGCGATGTCGAGCTATATCCAAACGCGAAT 540  
 Db 56808 TGTAAGGAATAAGAAAAGAGAAAGTGGACAATATCAATATATCTAAATCTGCT 56867  
 QY 541 TTGAAACCCCTCGAGAAATGACCGGACCGTGTGTTATGAGCTTCCTCAATATTTGTAATC 600  
 Db 56868 TTAATATGTTATGTAATAATGGGAGTCTTCAATATGAGCACAATAATGTTATGTAAT 56927  
 QY 601 GACGAGCGCGAACCCTGTCCTTACATACCTACACACCGTCTGTTCTTTTACTGGTTC 660  
 Db 56928 GATGAGCGGGACATATATCATCATATTTCTATCTGCTGTTCTTTTACTGGTTC 56987  
 QY 661 CTAAACAGCTGGTAAATACCCCTCTTTACCGGAGGCGGCGTTCGCTGCAATGATGTC 720  
 Db 56988 TATAATAGCTGGTTAACTCCCTCTATACAGAGCGGAAAGTTCATGTTATGCTGCT 57047  
 QY 721 GTCGGTTCGCAACGAGCAAAAGCGTTCAGTCAACTTACAAACGAGGACGCAAAAG 780  
 Db 57048 GTGGATCCCGACGAGACAGATGCTTCAGTCCATCTTATATCATATATACACAGAG 57107  
 QY 781 ACGAGATATCGTGTGCGAAGACATCTTAAATCATGATAGGGAAGAGTCTGATCT 840  
 Db 57108 CATCATCTCAGCATGTGCAATGTGTTATCCCTGCTGATCAACCAATAGATATGCT 57167  
 QY 841 GAGTACGTACATGAGAGAACTGGCGCTGTTTATAAACAAAGAGCGTGCACCGAT 900  
 Db 57168 GATTACATAGATCTAAATACAACTGGGCATTTGTTTATCAATACAGAGATGTAAGT 57227  
 QY 901 CTCAGTTCGACACCTCTTAAAAATTTTAGAGTATAATCTTCCCATCTCTGACGAAGTC 960  
 Db 57228 CCAGAGTTTGGACATATGCTCAAACTTTGGAATATAATCTCCCAATACCTGATGATAC 57287  
 QY 961 ATGAGTTAGTACAGATATGCTGTTCTTAAAGTAAAGTATGATGATGATGATGATGAT 1020  
 Db 57288 ATGAATTACATAGATATGTTTGTCCCAAGAGCAAGATCTCAATCTCTAGAGTAT 57347  
 QY 1021 ATTGGCTGGACCGGACTCTTTTGTACATAGCGAGTAAAGCGGTATCTAAACAACTTA 1080  
 Db 57348 ATTGGATGAGCAAGACTGTTTATCCCAATAGGTAAAGCGTACCTGCTCTCTTTA 57407  
 QY 1081 CACACATGTCTAACGTTAGGGGCGATACAGGAGACAGAGTCTTTACCTGTCGCGTG 1140  
 Db 57408 CATACTGCTGCTGCAGCTACTAAACAAACCGGGGCGAAGCTATTCACATGTCCCAAT 57467  
 QY 1141 GTGTGGAGGTGTTTGTGAAGCGGTTTGAAGGATACAAAGCGGCGCTCAACCTCAACAC 1200  
 Db 57468 GTGTGTGAATTTTCAAGCGGCTTTTCAAAATATAATCTCACTGAGCTGATGATCT 57527  
 QY 1201 CTCACGCTGACCGAGTGGGTGCAAAAAACCTCTTTAAGTTAAGTAAATTTTGGAGTTT 1260  
 Db 57528 CTAACCAACAGATGGTTAATAAGAACTTATTTAGATTAAGTAAATTTACTCAAAATTT 57587  
 QY 1261 GTGACCGAGCATGTCCATAGTCCGACCGGATCCACTGAAGCTTCGACCCGAGTTACC 1320  
 Db 57588 ATAGACCAAGACATGCAATGTGATCTACAGAAACATCTGAAGATAGTCTTAAGATTACA 57647  
 QY 1321 TTTATCACCAAGTTTGTAAAAACAGCCAGCTGCTTTTAAATGGAAAAACAAAAAATGTC 1380  
 Db 57648 TATGTCAAAAATATGACCGATAGCTATGTTCTCTCAATGGAAAAACAAAAAATCT 57707  
 QY 1381 ATATGGGTTTCAGGGTAGCTACTTTGAGTTCAAAAGAACTCTAGACAGCACTCTTC 1440  
 Db 57708 ACTTGTGGCTTTATGGGCCCATATGAATCTTTTAAAAAGGATATAGAATGTGAACATTC 57767  
 QY 1441 GTGAGACGATTTGCGCAAGACCGTCCGGAATACGTATACGTTTCTTAAACACGCTACTG 1500  
 Db 57768 ATAGACCCCATGACATGATCACTCACTGATTTTGTGTTTACAGTTTTTAAATATCTCTG 57827  
 QY 1501 TACAACCCCATGACTGTTTTCAGCGGTACGCGGTAGCTAGGTCAACGAGAAATATCTG 1560  
 Db 57828 TATAATGGGATGTAGTTTCTATATACGGGAGTTAAGGAAAAAACAATAGACTATCTT 57887

QY 1561 CAAGACCTAAATTTGACCCCTCCCG-----GCCCTCTGGCAACCGGCGGTGTA 1611  
 Db 57888 TCAAAACTAAATCAGATACCCATCCCTCAAATATCTCCACCCCGGAACCAAAAAAGT 57947  
 QY 1612 GACCTTCAACGGTCTGTAAGAGTTAAACCTGGAAGACACATCTTCTACCACTGTCG 1671  
 Db 57948 GATTACATATCTGATAGTCCAGAGAAACATTTGAGAGATAGATTTTACTATATGTCG 58007  
 QY 1672 AGTCCCGGCAACCGCGGTATACCTCCCTCCAGGTTTGTGTCGACAGTACTGCGCC 1731  
 Db 58008 CAGCCCGCAACCAAGCAATGTCACCTTTTACCTGTTTGTATGGCTATTACAATGCT 58067  
 QY 1732 CTAAGGACGCTGTTCCGCTCCAGAAATAAGGTGGCGTCTCGCTGGTTTGGCGGAGTTT 1791  
 Db 58068 CTAAGGAAATATTTTACTAAGCTTTCCTTGGCCATAGAAAGCTTTGGACATCTTTT 58127  
 QY 1792 GAGAAGAAACGTTTTCGCGTTTACGGTTAAATGCTGTAAGGAGCGAGTTGACTTT 1851  
 Db 58128 CTTCAAAAAAATTTGTAACATTTACAATAATATTTGATATCAGAGATGAGTTGAACCTC 58187  
 QY 1852 GTCTCCCTTCAGAACCTCTCAACGGCTGTTGGCGTTTGCATCGACCGTTGAAATCTAT 1911  
 Db 58188 ACATCTGATGCAACCCACATTTTGGCTCTTATGATATGATCTATCTGTGGAATCATAT 58247  
 QY 1912 AAATTAAGGGGTACACGTTTACCGGTAGCGTTCCGCTGCTGTCAAGGTTTGCCTACT 1971  
 Db 58248 AAGCTCAAGGATATATCAATCTCCCAATTTTGGAAAGATATGCTGTGTGGAGCTA 58307  
 QY 1972 AGTATGACCTCAGGAAGAGATGCTCTCCCTGCTGTGTCGAGGACTCTAGCGGTTTATC 2031  
 Db 58308 AGTAAGGATCTAAAAACCAAAATGCCCCAATATAGTAGTTGAAAGACTCCCTAGTTTCA 58367  
 QY 2032 CGGTGCTTAGAGATAACATAACCAAAATGACCGAAACCATGAGGAGCGGAGCGTTTTC 2091  
 Db 58368 TCTTGCTTAGAATATATGATAAATTTACAGAAACCCCTAGAGGATGGAGTGTGTTTC 58427  
 QY 2092 CAAGTGTGCTGTGGGGGACTATGGGTGAGTCAAAATTTGCCATGACCATCGTAAAG 2151  
 Db 58428 CATCTTTGAGTGGGGTGAATGAGTGTGAGCTCTAACTGGCTATGATTTGTAAAG 58487  
 QY 2152 GCACAGGATGCTGTTGGAGCGAGTAGCCGTAGTATTTGGGTCCCAAGAAACGTCAG 2211  
 Db 58488 GCACAGGATATCTTTGGACAAAGTTGCAATCTCAATTTGGMAATCATAAACATCAA 58547  
 QY 2212 ACAAGCCAGCTGATGATGAGCAATATCAAGGCTGTTAACTCAATTTTGGTCAATGAC 2271  
 Db 58548 CTTAGTCATGATATGTTGGCGATCTCGAGGGCCACGAGCCCGAGAAATGTTGGTAAATAG 58607  
 QY 2272 AGCAACCCCTTAAACCCCTCTCAGAGAACCCAGTCGATTAACACCTCCGCCAAGCATATA 2331  
 Db 58608 TGCACCCACATAAACAATATTTCCCTCCACGAGCCAAA---TTCATCTAGTCATATT 58664  
 QY 2332 GTCCGCGCTTCCACACCCCAACACACCCCTATCTACTAAATAAATAAATAAACAAG 2388  
 Db 58665 GTCAAGGCGCTCCATAACCCCAAAACCTCTGTTTATTAAATAGATGAGGAACAG 58721

RESULT 12  
 AF083424  
 LOCUS  
 DEFINITION  
 AF083424  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Ateline herpesvirus 3  
 Ateline herpesvirus 3  
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 Gamaherpesvirinae; Rhadinovirus.  
 REFERENCE  
 1 (bases 1 to 108409)  
 Albrecht,J.C.  
 Primary structure of the Herpesvirus ateles genome  
 J. Virol. 74 (2), 1033-1037 (2000)  
 MEDLINE 20091363





Qy	EMPPYINMAPDECIGKGPCKAKNGDLIIIOISCVFMOVGPDTTKMNLISLGTCAA									
	VEDSEVYEPSEADMLHGPTTILRDNVELIIGYINISNPDLYLDRATOIYNIKLSD									
Db	YRSKVGSIQVHMPDMGNGFMRSVSKISGIIAIDMLIVCKDKLSNRYLDTVA									
	DHCIGAKKEDSVKDIILPFGMSGPEGRKIGRYCVIDSLVLKLMFKEHYISEIA									
Qy	KNALITRVLVDPQGIIRVSCLLAAARAENILVSNVNTDQFGQATVPIIPFGY									
	KNAVLVDPFASVYSIOAHNLCSYTLIPHNSLHNPYLKSDSYETFLMSGPIHFVK									
Db	KHQTSLLKLTWLSKRAKROKLAQCEDDQTKTILDKQOLAIKVTCNAVYGTGV									
	ASGLLPCISIAETVLTQGTMLKSKAFIEEMTPVRLQBIIVPHKLHHEHAKRYVIG									
Qy	DTDSLVEVCVGVSDPVVAFEGDFLFTSEKLFNAPILKLESEKTLQCLLLAKKRYIG									
	ILNSDKLLMKGVLDVKTACKPVQKTSKILNLIILKDPREKAAQULLSKDDPDYAFRE									
Db	GUPDGKLLKVIDILNESYENLRGKVMDDLTSTELSRPISIKYENUPHLTVYKKII									
	KQOEPPQKHRIIPVYVGVKATKICSTMAEDFTYAVQHNIPITARDLYFDKLLHGVANI									
CDS	IOCLFSDSNKTSVLYNFVSIPLFSYELKSDHSTET"									
	20192..21412									
Qy	/notes="orf 10; similar to Raji LF1"									
	/codon_start=1									
Db	/protein_id="AAC95534.1"									
	/translation="MNPSTNFINGRWRTVTVSSGLFSISNNEIAPATPVARPVHLPV									
Qy	STBSIIDKRVSHGLVTLQFOQYNSVPAVSDTSEQPKPYVLPMTISDFINPLV									
	FVKGGIILIRKELGYIVIFPLRTDISIFPPDNDTDKTLALAGFRTVIEVENH									
Db	IRGPTKNTGSKVLMFRCEKIPLFVNIINFNKDSLELIFESLLIQEVLNTYIMTH									
	MCGSPEDQIEVQVHLNVINTTELVFHALPHELAPSGSILPIYLEADKILKPGSS									
Qy	VEISFSFVNRGLVDNQSALEFVASSNHTTKIVVPQIWIPIPLUSITVYNPGRNII									
	IKRGFCVAVAPCFYFLKAPQDCEDKTLDSQTSIHSWDLIIRKPGEGGTVVHAHL									
CDS	VLKEINFTEPMTF"									
	21415..22632									
Qy	/notes="orf 11; similar to Raji LF2"									
	/codon_start=1									
Db	/protein_id="AAC95535.1"									
	/translation="MATPPIHGRVWRGSMKQSRKFNFLFWSSSVHEHFTIITNQ									
Qy	EQIOLPDDVLLARCPISREIVGQKLPNFKFSRYPGTITTSFYVGHCENVIKIP									
	MTISDCDQELIFALKECIIPPGSKIFILPIPLDKLDGLYLCLBDETSIMTTS									
Db	CMQMGVLTSETPOVFLKGSVLTKNELPHLMAOKTPEFNKMAVHTVONEVECN									
	SIYKGDHRLALOKKSDVDNPFEPVILGLTMNKALJAFQHNPPYFCFMDWGRSIP									
Qy	IIVYGSIRIPDRHVAVKXNMYSYANITAMINCHETCPDHLSCWEKPERPVY									
	INTVNTINAITISSGTLKGEAVFLAPFLCKKIISKKHVLTLSAVTLPGNVITS									
CDS	NKLPKLDLITTK"									
	complement (23207..24028)									
Qy	Query Match									
	Best Local Similarity 9.0%; Score 812.8; DB 14; Length 108409;									
Db	Matches 1477; Conservative 0; Mismatches 997; Indels 16; Gaps 5;									
	1 ATGAGAGTTTCCGCGATGACCC---AAACAGTCCGAACCAAAATCCGGGGTTTCATCTTG 57									
Qy	57799 ACGACAGTTTGTGGATGAACCGGAATAATGGCAGAACTAAGTCCAGAAATTCATCTG 57858									
	58 AACATGAGTCCGATGCCAAGTCAGGGTGTCTGATCATCATGTCAGTCCCTGTCAAT 117									
Db	57859 AACATGAGTCTGATGCAAAAGTCAGAAATTCATGTCGAGACTATAAGAAAGCTGTCAAT 57918									
	118 ATAACATACAGCCACCGGAATGGGTGGTACGACCTGGCCCTTCGATCCGCTGAAGAC 177									
Qy	57919 ATTACGACTCGCTCCAGAAATGAATTTGTAATAATAATCAATTTGACCCAGAAAATGC 57978									
	178 TCGGGCCGTTCTTGGGTTTACGGTTTATCTAATTTACGGGAACCTGCTGCTGGAAA 237									
Db	57979 TCTGGACCTTGTCTCCATTTACAGTTATGTAAATTTACTGGCACTCGCGGGCTGGAAA 58038									
	238 AGTACCAAGCATATCGGCCCTGTACCAAAATTTAAACTGCTGATCATCGGCGCGACCC 297									
Qy	58039 AGTACAAGTATCTGACCCCTACATCAAAACCTTAACTGCTGATTAATCTGGCGCAACT 58098									
	298 ATAGCCGCAAGAACCTATCGCTGCGCTTAAAGAGCTTCTGTCCCAAGATCTTCAGCGCT 357									
Db	58099 GTTGCTTCTCAAACTTATCCAGTGTCTGAAAACATATCTGCCCAACAATTTATAATGCA 58158									
	358 TTTGGCTTTAAGGCGGACATCAATATAGCCGTCAGAAAAGCTCATCAGACCGGAGCC 417									
Qy	58159 TTTGGGTTCAAGAGCAAGCAATTAATTTCTACCCAGATCTGTGCTCGCACAACTATG 58218									

Qy	418	GTATCCATAGAGCAAAATTCAGCAACAGAGAGCTATCGAAGTATTCGGCGGTATATAGTGGAC	477
Db	58219	GATACAAITGAACAAATCCAAAATTTTGAGTTATGTAATATTTGCCCATATTGTCRAGC	58278
Qy	478	ATTATGAAGAGGTTATGGCGAAAAAACCACATGGCAATGACGGGATATATCAACGCG	537
Db	58279	ATTATTCAGAGCTTTTCTAAAAAGAGAAATTTAGGACAGTACTCAAGCAITTAGCTTATCA	58338
Qy	538	AATTTTGAACCCCTCTCGAGAAATGACCGGACCGGTATTATGACCTTCCAATATTATTGTA	597
Db	58339	GCTTTTAATTTGGCTATCTAAATGAC---AACGGCCCTTGTGGACCACTAAACATCATTTGTC	58395
Qy	598	ATCAGCAGGCGCGGAAACCCCTGCTCTTTACATATCTACCAACCGCTGTTCTTTTACTGG	657
Db	58396	ATAGACGAAGCTGGCACAATGCTCTCATATATCTCACAGCTGTGTGTTTGTATTG	58455
Qy	658	TTCTTAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGGGGTTCCTGTCATAGTA	717
Db	58456	TTTTTTAATAGTTGGTTGAACACTCCCTTATACAGGCTCTGGAGCGTTCCATGATGTT	58515
Qy	718	TGGCTCGGTTCCGCAACGACAGACAAACGCGTTTCAGTCAACTTCAACACCGGACGCA	777
Db	58516	TGTGTAGGATCGCTTACAAACAGATGCATTCATATATACCTATATACCTCGACGAA	58575
Qy	778	AAGACGGAGATATCGTGGTGGAGAAACATCTTAACATTCATGATAGGGAAGAGGTGCTA	837
Db	58576	AAATACAGTGTAAATGGAATGCGATAACATCTGTCTTTCTCATAGAAAAAATAAGTTGTG	58635
Qy	838	TCTGAGTACGTACACTTGGAGAGAACTGGGCGCTGTTTATAAAACAACAGCGTGCACC	897
Db	58636	GCTGAGTACATCTCCATTAATTAATCTGGGCTTTATTATAAAACAATAAAGGTGTACA	58695
Qy	898	GATCTGCACTTCGGACACCTTCTAAAAATTTAGAGTATAATCTTCCCATTTCTTCGACGA	957
Db	58696	GACCTGAAATTTGGACATTTACTGAAAAACACTTGAATATAGCTTAAAAATTTCTCTTAA	58755
Qy	958	GTCATGAGTACGTAGACAGATTTGCTGCTTCTTAAAGTATAGATTATGGATCCTTTAGAA	1017
Db	58756	ACAATAGATGATGTAGACAGATTTGCTGTACTTAAAGCTCAGATTTCTCAATCTCAATAGAG	58815
Qy	1018	TACATTTGCTGGACCCGACTCTTTTGTTCATAGCAGGTPAAAGGGGTATCTTAACAAAC	1077
Db	58816	TTTCTAGSATGACACAGACTCTTCTGTCTCATGCTGAAGTAAAGAGCTATTTGACATCA	58875
Qy	1078	CTACACATGCTTAAACGCTAGGGGGCGATACCAGGGACAGAAAGCTTTTACCTGTGCC	1137
Db	58876	CTTCATACAGACTTAACTCTGGAATAACTTATCATGAGGCAAGAGCTTTTACTTTGCTCT	58935
Qy	1138	GTGGTGTGCGAGGTGTTTGTGAAGCCGTTTGAGGAATACAAAGCGGGCGCTCAACCTCACC	1197
Db	58936	ATTGTTTGTGAAGTATTCACCTCCAGCATTTAAATGAGTATAAGTCTCATGTTTAACTACCT	58995
Qy	1198	AACCTCACCGTGACCGAGTGGGTGACAAAACCTCTTTTAAAGTTAAGTAATATTTCGCGAG	1257
Db	58996	TCTTTGACAGCTACAGAAATGGTTATCAAAAAATTTACACCGATTAACTATTTACAG	59055
Qy	1258	TTTGTGGACACGAGCATGTCTCATAGTCGCAACGAATCCACTGAACGTTCCACCCAGGTT	1317
Db	59056	TTTATAGCCAAAGATGTTGCGGCCACCATACAGAAACTACAGACATCAACAAAAAGTA	59115
Qy	1318	ACCTTTATCACCAAGTTTGTTHAAAACAGCCAGCTGCTTTTAAATGGAATAACAAAAAA	1377
Db	59116	ACATTTCTTTACAAAATATGTAATAAACAACATACATTTCAATTAATGAAAAACAAAAAG	59175
Qy	1378	TGCATATGTGGGTTTCAGGGTACGTTCTTTGAGTTCAAAAGAAATCCCTAGACAGCGAACTC	1437
Db	59176	TGTGTGTGTGGTTATGCGAGGAACCTTAAACATTTTAAAAAAATATTATAGAAAGTGAAGCT	59235
Qy	1438	TTGCTGGAGACGGATTCGCAAGACCGCTCCGGAAATACGTATACGGTTTCTTAAAACACGCTA	1497
Db	59236	TTTATASATAGCCACACAAATGACCAACCTGAGTTTGTTTACAGTTTTTTGTGTACAATA	59295
Qy	1498	CTCTACAACGCCATGATCTCGTTTTCACGCGTACGGCGTGACTAGGTACACGAGAAATAT	1557

Db 59296 ATTTATAATTCCTATATATATTTTTCATAATTTATGGCGTTCTTCAAAATAATGAGGCTTAT 59355  
Qy 1558 CTCGACAGACTAAATTTGCAACCCCTCCGGCGGCTCTGGCA-----ACCGGGCTGTA 1611  
Db 59356 TTGCATGATCTAGCAAAATTTAAAGCTTACCAGAAAATTTTAAACATCTGTACACAAGCA 59415  
Qy 1612 GACCTTCAAAACGGTTCCTGAAGAGTTTAAACCTCGAAGACGACATCTTCTACCACTGTGC 1671  
Db 59416 GATTTAGATATTGAAAAGGAACACTCATGCTAGAGACGATGTGTTTATCATCTGTT 59475  
Qy 1672 AGTCCCGCGCACCCCGGGGTATCACTCTCCCTCAGGTTTGGTTCGACACGACTACTGCGCC 1731  
Db 59476 GTTCTCTCCACCCACAGCTTCATCCGATCTCTTCCATGCTCTTAATAAATGGTACACTGCA 59535  
Qy 1732 CTAAGAGAGCTGTTCGCCCTCCAGATAAAGGTGGCGTGTGCGTGGTTTGGCGGGAGTTT 1791  
Db 59536 CTAAGAGACATATTTATATCTCTGCTAAAGTTGGCAACAGCTGTTGGTGTGAGACATTT 59595  
Qy 1792 GAGAAGAAACGTTTTCGCGGTTTACGGTTTAAATCATGCTGTAAAGGACGAGTGAATTT 1851  
Db 59596 TTAGACCATGAGTTTCATCATTTTACAATTAACATGTTAGTAGATTAATAGAAATTT 59655  
Qy 1852 GTCTCCCTTCAGACGCTTCACGGGCTGTGGCGTTTGGCATCGACCGTTCGAATCGTAT 1911  
Db 59656 ACCTCTACCAATGACGACTTCATGACTTTTGAATATGCTTCAACTGCTCGAATCTTAC 59715  
Qy 1912 AAAATTAAGGGGTACAGTTTTCACGGTAGCTTCGGTGTGCTGAGGTTTGGCACTTC 1971  
Db 59716 AAATTACAGGTTATATCTTTTACCAGTTCATTTTGAAGAGCTCAGACAACGCTTATA 59775  
Qy 1972 AGTGATGACCTCAGAGAGATGCTCCCTCGTGTGCTGACGAGCTAGAGGCTTTTATC 2031  
Db 59776 AGCAAGATTTTCAAGATAAATGCCAAATTTATAGTGAAGATTTAGCGGATTTATA 59835  
Qy 2032 GCCTGCTCAGAGATTAACATAACAAATTCACGAAACCATGAGGACGCGGCTTTTC 2091  
Db 59836 GCATGTTTGAAGAAAATGTTAAACAAATGATTGAAAATTTAGATGATGAAAAGCTTT 59895  
Qy 2092 CAAGTGTCTGCGGGGACTATGGGTGAGCTCAAAATTTAGCCATGACCATCGTAAAG 2151  
Db 59896 CACCTCTGTAGCCCGGAGATTATGGCATAAGCTCCAAACTTCCCATGACTATTGTTAAG 59955  
Qy 2152 GCACAGGAATGCTGTGAGCAGTAGGCGTAGTATTTGGGTGCCACAGAACGCTCCAG 2211  
Db 59956 GCCAAGGTATATCTTTGATAAGTTGCTATATGTTTGGAAACCATAAACCTGAAA 60015  
Qy 2212 ACAAGCACGTTATGTAGCAATATCAAGGGCTGTAACTCAAAATTTTGGTCACTGGAC 2271  
Db 60016 ATAGTCACATTTATGTTGCAATTTCCAGAGCAACCAACCCCAACCACTTGTATGGAC 60075  
Qy 2272 AGCAACCCCTTAAACCCCTCCTCAGAGAACCACTGATGATAACACCTCCGCGCAAGCATATA 2331  
Db 60076 CACAAACCTTTAAAA---CTAATGTGTAATCATGCCCCCATGTGTGCTTCCAGCACATA 60132  
Qy 2332 GTCGCGCCCTCCACAAACCCCAACCAACCCCTCATCTACTTAATAATAAATAAACAAGCAA 2391  
Db 60133 ATCAAGGCACCTTAACAATCCCAACACCCCTATTAGTATATTGAATAAATAAATAAAGTT 60192  
Qy 2392 ATGAAAACATTTGTTTTTTTATTCAGTCCCAACCAACCGCCAGGATGTTGTCATTTCCAC 2451  
Db 60193 AAATTAATTTAAAGTGTAT-TTTAATAAGCACCGGCCAAACATAACAATTAATGCCCG 60251  
Qy 2452 ACACCGGGGGGTGGCGCCCATAGTTTTCAC 2481  
Db 60252 TAATTGGGGAGTTCCAGTCATTTTGGTC 60281

RESULT 13  
HSGEND 112930 bp DNA linear VRL 25-OCT-2002  
LOCUS Saimirine herpesvirus 2 complete genome.  
DEFINITION X64346  
ACCESSION

X64346.1 GI:60320  
VERSION  
KEYWORDS alkaline exonuclease; complement control protein homologue; complete genome; cyclin homologue; DHFR gene; dihydrofolate reductase; DNA polymerase; dutPase; G protein-coupled receptor homologue; glycoprotein; glycoprotein B; glycoprotein K; helicase; HVS-R transactivator; Ig-G gene; integral membrane protein; major capsid protein; major ssDNA-binding protein; mCpH gene; ribonucleotide reductase, large subunit; ribonucleotide reductase, small subunit; saimiri transformation-associated protein; sCpH gene; STP-A11 gene; thymidine kinase; thymidylate synthase; TK gene; transport protein; tyrosine-specific protein kinase; uracil DNA glucosidase.  
SOURCE Saimirine herpesvirus 2  
ORGANISM Saimirine herpesvirus 2  
VIRUSES; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.  
REFERENCE 1 (bases 1 to 112930)  
AUTHORS Albrecht J.C., Nicholas, J., Biller, D., Cameron, K.R., Biesinger, B., Newman, C., Wittmann, S., Craxton, M.A., Coleman, H., Fieckenstein, B., and Honess, R.W.  
TITLE Primary structure of the herpesvirus saimiri genome  
JOURNAL J. Virol. 66 (8), 5047-5058 (1992)  
MEDLINE 92333688  
PUBMED 1321287  
REFERENCE 2 (bases 1 to 112930)  
AUTHORS Albrecht, J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JAN-1992) J. Albrecht, Inst. fuer Klinische und Molekulare Virologie, Friedrich-Alexander Uni, Loeschgestrasse 7, D-8520 Erlangen, FRG  
FEATURES  
source  
1. 112930  
/organism="Saimirine herpesvirus 2"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10381"  
complement (220. 714)  
/gene="STP-A11"  
complement (220. 714)  
/gene="STP-A11"  
/note="ORF 01"  
/codon\_start=1  
/evidence=experimental  
/product="saimiri transformation-associated protein"  
/protein\_id="CAA45623.1"  
/db\_xref="GI:60321"  
/db\_xref="GOA:P18347"  
/db\_xref="SWISS-PROT:P18347"  
translation="MARGLEGDPQENDESNDPPHNTDERSDGDGPTPYLPVTLN  
AGFGPNPFCLLHGVQESGEPRTALSAVGLPTSGRSSSHLSTPVGLSAVR  
SGCGAGSEEHVYAEVSLSEHREQDKCTCSVTILLLLVILLLIILGLMLVIM  
FKW"  
complement (3969. 4532)  
/gene="DHFR"  
complement (3969. 4532)  
/gene="DHFR"  
/note="ORF 02"  
/codon\_start=1  
/product="dihydrofolate reductase"  
/protein\_id="CAA45624.1"  
/db\_xref="GI:60322"  
/db\_xref="GOA:P09503"  
/db\_xref="SWISS-PROT:P09503"  
translation="MVQALNCIVAVAGNMIGKQGNLPWRLMDFKHQRMVTTSSV  
PDKNLYMGKTFSTPEKNRPLKGRINVLSKELKELPHRAHFLAKSLDDALKL  
QPELANKVDVWVIGSGSVYKAMSYPGDLKLFVTRLMQDFECTFPFEDLEKYL  
IEYPSVLSNVQEEKSIKYFEVIEKNH"  
6821. 10561  
/note="unnamed protein product; ORF 03  
similarity to ORF 75 and EBV BNRF1"  
/codon\_start=1  
/protein\_id="CAA45625.1"  
/db\_xref="GI:60323"  
/db\_xref="GOA:Q01000"

```

/db_xref="SWISS-PROT:Q01000"
/translation="MLYLHNLNLEAQLPHERLATSVFLEHGLQVTSFNPFTTTGTVY
VYNVSNASLEPVLKFLQSPVSKPFFXSEQTLSFSFGHLOQSQTTFSGKLDVY
VKMLSTPOEDNFAETVEYRTHIICDLQPKGRHIRHLOMLCNPTVGLTNF
VIPDKELEEDNABSAVASAPLSDPLMTWSNGLNGLGVSSOYVFKGAAPKVE
KNSELAISMDIDYKNPWSFMLYDGLWHAIEKGLIYSGSFPTPSLTSVKNWACHY
VNPESACQIHYMPLQTHLSLITGSLPKMLCAHLNLLKSSMHHIOGFRLVPKQ
LSHLPOQIKYNGFISLQTLINMANEHPLOQKTYLQGLNFFTHQKFLTCQBSG
KEIGSLATLAFSPNSVKTYPGITGVNMTSLPVSAMKAAVACVOCGARLVLSALP
POVYKALTPSPNSNRNKKQYFENLITSLIIAINKNVAIDAAQLTKACTYAGG
RFKIGILLTLEHPSGVKDETPYPIRLPDRFKPFPILLSPTSLFLEERINMDC
NLDRITKILHEHPSGVKEFTVNHDTKLISGRVATSIIVGPWOLPVSIVLPMHPC
TTESVRPNWDYETDIDSHCVFEAAVAGISALGSESTLITQADLAKVGTIRATEIL
NLSVPRNIGNIVQSLITPYATHAVSTYLQVLMETAKTFCALSVSTFWANATEG
GASIVASAI VNLIDSVSLITPKFNNSVFLFLLTETDLSFGVAQQILEKTFIGEI
PSATSPVMLKMSLVLQTLIKDSVSVSGDGLVATVAEMALSGGKGRVYVPHG
EDATKFTCSLTPGVVIVSQGSXMYIVQQLHSENINFIQIGESTLSTFSISQNLTK
VHEPLDFKSAWRGFSDACBPQCIHPRPYRMQIATVPKCPVGPCRFHTVIVYLLPN
SVPHGLNATAEAGFQPRLYSIHQPSKTVNPDTPVWGFFIVGASNVODEDVGMRAL
IAQLKSHVAFORDLRLMLAKPDVFSVAIGALACELLYNKAIGYNKPSDITYMTCVKN
SRKTESWNIYIPESTKATAFQSLKNSLIPICWTQGTHLKFFHPKPMLEKMEBSGNVS
SMFYGHSLSGSFPAQNYELTPNGENAIAGVGSADRHALLHDFPCLNMLQWYVPELE
NPPLKVPKMTFLDLHKWQITVGASPPESRSDPLRSFV"
join(10912, 11773, 11774, 11966, 11967, 12013)
/ gene
CDS
/ gene="sCCPH"
join(10912, 11773, 11967, 12013)
/ gene="sCCPH"
/ note="ORF 04b"
/ codon_start=1
/ evidence=experimental
/ product="complement control protein homologue"
/ protein_id="CAA45627.1"
/ db_xref="GI:60325"
/ db_xref="GOA:Q01016"
/ db_xref="SWISS-PROT:Q01016"
/ translation="MYTLHYICLVLCVIVFVWTLSCPTNRQVSVKYNLNTYSGPY
NGTILVHTCREGYAKRPQVQVTCVNGNWTVPKCKCKKCTPQDLNGRYVTGNLY
YGSVITYTCNSGYLSIGTTSACLLKGRGVDMTPRPPICDIKKXPPQIANGTHTN
KDFVYLDYTVTSCNDETTLTLTPSSKLCSETSGSVNGETKCSFIFCKLQPVANA
YVEVKRNTSNQYLIHNKCYKGMPLYGTEPNTCNHGVWSPAIPCWMKISSPKGDMPG
INSNEDNSTPGSACGACGSGSNYPSS"
10912..11994
/ gene="mCCPH"
10912..11994
/ gene="mCCPH"
/ note="ORF 04a"
/ codon_start=1
/ evidence=experimental
/ product="complement control protein homologue"
/ protein_id="CAA45626.1"
/ db_xref="GI:60324"
/ db_xref="GOA:Q01016"
/ db_xref="SWISS-PROT:Q01016"
/ translation="MYTLHYICLVLCVIVFVWTLSCPTNRQVSVKYNLNTYSGPY
NGTILVHTCREGYAKRPQVQVTCVNGNWTVPKCKCKKCTPQDLNGRYVTGNLY
YGSVITYTCNSGYLSIGTTSACLLKGRGVDMTPRPPICDIKKXPPQIANGTHTN
KDFVYLDYTVTSCNDETTLTLTPSSKLCSETSGSVNGETKCSFIFCKLQPVANA
YVEVKRNTSNQYLIHNKCYKGMPLYGTEPNTCNHGVWSPAIPCWMKISSPKGDMPG
INSNEDNSTPGSACGACGSGSNYPSS"
10912..11773
/ gene="sCCPH"
/ number=1
11774..11966
/ gene="sCCPH"
/ number=1
11967..12013
/ gene="sCCPH"
/ number=2
complement(11995..12285)
/ note="unnamed protein product; ORF 05
KFLF1"
/ codon_start=1

```

/protein\_id="CAA45628.1"  
 /db\_xref="GI:60326"  
 /db\_xref="SWISS-PROT:P24912"  
 /translation="MTFIFLHKAKHKLQAENSYNDPVVEIEIQLSEECNRNNT  
 SLNDSIYDDVCYPEDDDYDGGCAYTELDASRVYGLQRTCNDSHYIEEIG"  
 12584..15970  
 /note="ORF 06  
 KFRP1"  
 /codon\_start=1  
 /product="major ssDNA-binding protein"  
 /protein\_id="CAA45629.1"  
 /db\_xref="GI:60327"  
 /db\_xref="GOA:P24910"  
 /db\_xref="SWISS-PROT:P24910"  
 /translation="MAATKAQPSADNLGRAPVEPCGYIYVYKPEGPPFKESALLGNK  
 NYGASMLPLLSLTVESNEFSNVKAVHKIMDTLLVRVAVHRAEIVFENTDLPE  
 PI FVPGGLDILCSARSLEFGYTNFVPRITDLRDTVDIKDYAPYSEDSCEMAVYVTEG  
 FKERYFNGVPIIQAQKKVOINGREAVKIPLYDEDFSKSHEHLPRFYIPSVSKYLH  
 DSVFTSAQRILRDESVIVRASQSIQDYIKLAVIKNSKDFSLQSVKQCDASAFVH  
 IDCIAELNAGYSLFIEAPQDPCCAVLDYTSWPIFETAETEDRIKAIQDNWAMSVH  
 VYHLFSTNSVYLITKINKQTSNKRQNTVYFMQGLAYAADATQRENGEPFAGS  
 VYVTFSGYVTHLALASSFPHLLARNCYMQFCQKSTNNAYSPVQVGTAAAS  
 DLKSCGTGTCASCIHTLVELKDRFPVLGSGRRDPVYVTVGVGVQVNDMLGNFAT  
 PREKDEAVQNAESEKTYMQLIQNVVEKLSMGVTEGTVGSELIITDIQSFLTFPDI  
 DNVDSVEVYKMNCLKNNINFRITIKIVHVLHYCCNVFQWQAPCAMFLNLFYKSVLA  
 IITQICLPYVITQDDNPSGMESEWLKVHYQITWNFKESCDLRGVLIVGSHRKIKH  
 TDMFCDLNDSALSGIIVPMKQVRILAKALLTPVKTIKIKRIVFNSNMTETIQSG  
 FIKAKTQDSIVYGPYKMLSHKVMFPNAKISALYMLKHSQKQLFVLPQISSE  
 NMVELANVYESSKMDMNVLDITPTLLTYAKRVNLNLTITCGQTQVATLPLOCL  
 LPTLOTISATPEVHLLDQMSVDHYLSSIKDKHALTVOTLKEDTATVQKQPIVLT  
 VPLVNYKTGINGNTQIFQCGNLGYFMGRGVDRNLIPDSTGFRQNNSSYMRHRHFM  
 TPMVAHLVKGNLNLITFEVIRKQVNIPEDKDNLNIPDNVLVBLVKGLGDCSEN  
 ITEDDLQYLGEIYIMSDIEWSRFQITDGPAPWSVENVTKVLGCNKQKEECKEFVGV  
 BEQISCVPVQPEFAPQATLSTLAASRRKRIITSLSDIDL"  
 15977..18016  
 /note="ORF 07  
 KFRP2"  
 /codon\_start=1  
 /product="transport protein"  
 /protein\_id="CAA45630.1"  
 /db\_xref="GI:60328"  
 /db\_xref="GOA:P24911"  
 /db\_xref="SWISS-PROT:P24911"  
 /translation="MACHLAAYVQIYGLTLDVSLITFVDPAHNTWKTVTKNTEKINK  
 TYLOIMPLLNKNNITESLVAELQHLHLNKLIVLETFSAHLSDDYNNVFENHLSLAP  
 CSRKHSIVFQYNNCSYKNCIINIEFSKRLSSVYFCIRSCDALRGINHVIDFLG  
 HLRGVSPILPDTTSLNIPCYCLNEHMLPNQGESIPSLMMCVNKHVCKQLANPEPI  
 QCMFENLLQRIHIVENSKQEQNTCSIDNQIRDAASDKUNQNTIFENISAPYELIS  
 NLITYMSSGAKHKKAVENSMKLLSYEAKQNSYKIKCNSTHFFDEKAPYPIES  
 PCGGIFSSVDVDTVSLKSDCSLAFKQVYQOOLIKQNELFVRINKLQEGEDVSHAA  
 SAVPLSDKATVTVDOVLHDAHAKDVALQVTKDGLKSLYTCITQAGVALNLSLRK

	Query Match	8.5%;	Score 770.4;	DB 14;	Length 112930;
	Best Local Similarity	58.3%;	Pred. No. 6.9e-192;		
	Matches 1429;	Conservative	0;	Mismatches 1006;	Indels 15; Gaps 4;
Qy	34	GAACCAAAATCCGGGGCTTCATCTTGAACATGACGTCGCGATGCCAAAGTCAGGGGTGTGCTG 93			
Db	63007	GAATTAAGCCCTGAATTCATCTGAATGATGACTTCAGATGCAAAAGTCGGAATATTGTGA 63066			
Qy	94	GATCAGCTCAGTCGGCTGTCTCAAAATATACTACCAAGCCACCGGAATCCGGTTGGTACGAC 153			
Db	63067	GAGAAGATTAGGAAGCTGTCAAAATATTACTACTGTCTCTCCAGAAATGACTTTGTATAT 63126			
Qy	154	CTGGCCCTTCGATCCGGCTGAAGACTCCGGGGCGGTTCTTTCGGCTTACCGTTTATCTAATT 213			
Db	63127	GACCAATTTGACCCAGAACAAATGCCCCAGAGACTCTGCTTCCATTTTACTTGTATGTATA 63186			
Qy	214	ACGGGAACCTGCTGGTGTGGGAAAAGTACACGATATCGGCCCTGTACCAAAATTTAAAC 273			
Db	63187	ACNGHAACCTGACGAGCTGGGNAAGATACAGCATTTTCAGCTCTTTACCAAAATCTCAAT 63246			
Qy	274	TGCTGTATCAGCGGGCGGACCAACCAATGCGGACAGAAACCTATCCGCTCGCTAAAGACG 333			

63247	Db	TTCTCTATTACTGTTGCAACAACAGTAGTACTCTCAAAATCTATCAAGATGCTTGAAACA	63306
334	QY	TTCTGTCGCCAGATCTTCAGCGCTTTGGCTTTAAGAGCGGACACATCATATATAGCCGTC	393
63307	Db	TACTGCCCTACAATTTTTTAATGCAATTTGGATTCAGAGTAGACATTAATTTCTCCCA	63366
394	QY	AGAAAGAGTCTATCAGACCGGAGCGGTATCCATAGAGCAAAATTCAGCAACAGGAGCTATCG	453
63367	Db	AGATCTGTGCTCGCAGAACATTAGACTATTGAAACAATACAAAATTTTGAATTTATGT	63426
454	QY	AGATTTGGCCGGTTATAGTGGAATATATGAAGAGGTTATGGCGGAAAAACCAATGGC	513
63427	Db	AAATATTTGGCCAACTACTGACAGCAATATCAAGAAATTTTCTAAAAGAGAAATCTTAGGC	63486
514	QY	ATGTACGGAGCTATATCCAAACGCGAAATTTTGAACCCCTCTCGAGAATGACCGGACCGTCT	573
63487	Db	CAATACTCAAGCATCAGCTTAGCAGCTTTTAAACATGCTAGCAAAATGAC--AACTACT	63543
574	QY	TTATGAGCTTCCAATATTTATTTGTAATCGACGAGCGCGAACCTGTCTCTTACATCTAC	633
63544	Db	TTGTGACACAAATGTAATTTGTTATAGACGAAGCTGGACCTGTCTCTTCACATCTTA	63603
634	QY	ACCACCGTCGTCTCTTTTACTGGTTCTCTAAACAGCTGGCTAAATACCCCTCTTTACCGC	693
63604	Db	ACAGAGTTGTGTTTGTATTTGTTTATTAACAGCTGGCTTAACACCCCTTATACAGG	63663
694	QY	CAGGGGCGGTTCCGTGTCATAGTATGCTGCGTTGGCCAAAGCAGACAACCGCTTCCAG	753
63664	Db	TCTGGGCGAGTTCCATGCAATGCTGTGTGTCGGTCTCTACTCAAACTGATGCAATTCAT	63723
754	QY	TCAACTTACAACACGGGACGCCAAAGACGGAGATATCGTGTGCGAGAGACATCTTAACA	813
63724	Db	TCTACATACAATCACATACAGCAAAAGTATTAACATATGAGAGTGACAAACATATTATCA	63783
814	QY	TTCATGATAGGAAGAAGTCTATCTCAGTACGCTACACTTTGGAGAGAAACTGGGCGCTG	873
63784	Db	TTTATTTTGGAAACAAAGTTGTCTCTGAGTACATCTCCCTCACTAATTTGGGCTTTA	63843
874	QY	TTTATAAACAAACAGCGCTGACCGATCTGCAAGTTCGAGACCTTCTTAAATTTTAGAG	933
63844	Db	TTCATAAACAAATAAAGTGTTACTACCCCTGAAATTTGGTCACTTACTTAAACACCTTGAA	63903
934	QY	TATAATCTTCCCATCTCCTGACCAAGTCACTAGTTTACGTAGACAGATTTGTGTTCTTAA	993
63904	Db	TATAGCTTTAAATTTCTCTCTAAACAATGGAATATATAGATAGATTTGTTGTACCAAAA	63963
994	QY	AGTAAATATGATGCTTTAGATATACATTTGCTGGACCGCACTCTTTTGTCTACATAGC	1053
63964	Db	GCTCAGATTTCTAAACCAATGAGTTTCTAGCTGGAACAGACTATTTTGTCTCTGCT	64023
1054	QY	GAGGTAAAGCGTATCTAAACAAACCTACACACATCTCTAACGCTAGGGGCGATACAGG	1113
64024	Db	GAAGTTTAAAGCTACTTATCTTCACTTCTACTGCAATTTAGTTTACTGGAACTAATGTATCA	64083
1114	QY	GACACGAGCTCTTACCTGTCCTGCTGGTGTGCGAGGTGTTTGTGAAGCGCTTTGAGGAA	1173
64084	Db	GGGGCAAGTTGTTTACTTTGCCCTATTTGTTGTGAAGTATTTCACTAAAGCTTTTAAATGAA	64143
1174	QY	TACAAGCGGCGCTCAACCTCAACCACTCACCGTACCGTGACGAGTGGGTGACAAAAACCTC	1233
64144	Db	TATTAATCTCATGTTTAACTTACTTCTTCTTTTGACAGCTACTGTAATGGTTATCAAAAAATTA	64203
1234	QY	TTTTAGTTAAGTAAATTTTTCGAGTTTGTGGACCAAGGACATGTCTCATGTGCGCACGGAA	1293
64204	Db	CACCGTTTAAAGCAACTATTTCCAGTTTATAGACCAAGATATGACTGTCTATTCTPACTGAA	64263
1294	QY	TCCACTGAAAGTTTCGACCCAGGTACTTCTTATACCAAGTTTGTATAAAACAGCCAGTG	1353
64264	Db	ACTACAGACAGCTCAACAAAGTAACATATCTTACAATATGTTTAAACACACATATTT	64323
1354	QY	TCCTTAAATGGAACCAACAAAAATGCATATATGTTGGTTTTCAGGGTACGTACTTTGAGTTC	1413
64324	Db	TTCTCTCAACCGGAAAACTTAAAGATGTGTGTGTATGTAGGCATTTATAAAATTTT	64383
1414	QY	AAAAGAACTCTAGACAGCAACTCTTTCTGTGAGAGCGATTCGCAAGACCGTCCGGAATAC	1473
64384	Db	ABAAAATACTAGAAAGTGAATCTTTTATAGATAGTCAATGCTAAACGACCACTGAAATTT	64443
1474	QY	GTATACGGTTTCTTAAACACGCTACTGTACACGCCATGTACTCGTTTCAACGCTACGGC	1533
64444	Db	GTTTACAGTTTTTTTGTGTACAATCTTATATAATTTCTTTATATAATTTTCACTAATGCG	64503
1534	QY	GTGACTAGGTTCACACAGAGAAATATCTGCAAGAACCTTAAATTTTGCACCCCTCCCGCGCT	1593
64504	Db	GTGACTCAGAAAAATTCAGTCTTCTTAAATGACTTTAGCAAAATTTTAAAGTTTACCAGAAAT	64563
1594	QY	CTGGCAACCGGGCGGTCT-----AGACCTTCAACCGTTCTGTGAAGAGTTAAACCTGGAA	1647
64564	Db	CTAACCACTTGTATACACAAACAGACTTTAGATATTTGAAGAGAGAGCACTCATGTTAGAA	64623
1648	QY	GAGCACTCTTACACAGCTGTGACGTCCCGCCGACCGCGGGTATCACCTCCCTCCAG	1707
64624	Db	GACGATGTGTTTATCATATGTTCTCTCTCCACCACTCTGCTCATCAGCTTCTCTACCT	64683
1708	QY	GTTTTGTGTGACACGTACTGTGCGCTTAAAGACGTTGTTCGCTCCAGAAATAAGGTGGCG	1767
64684	Db	TGCTTAAATAGCTTGTATACTGCACCTGAAAGACATATTTATATCTCGCTCTAAAAATGCA	64743
1768	QY	TGTCGCTGTGTTTCGCGGGAGTTTGAAGAGGAAAAGTTTTCGCGCTTTACGGTTAAACATG	1827
64744	Db	ACAACTTGTGTTTATTAATTAATTTTATAGATAGAGAGTTTCACTCATTTAATATG	64803
1828	QY	GTGTAAGGACGAGAGTTGACTTTGTCTCCCTTCAGAACGTCCTCAACGGGCTGTGGCG	1887
64804	Db	CTAGTCAGAGATAATATAGAAATTTACATCTACTAATGGAAGACTACATGGACTCTTAGAA	64863
1888	QY	TTTGCACTGACCGCTTGATCGTATAAATTAAGGGTACACGTTTTTACGGTAGCGTTC	1947
64864	Db	TATGCTTCTACAGTTGGAATCATACAAATTAAGGCTATATCTTTTACCAGTCACTTT	64923
1948	QY	GGTCGCTGTGAGGGTTTGCACCTCAGTGATGACCTCAGGAAGAAGATGCCCTCCCTGGTC	2007
64924	Db	GGAAGATCTCAAACTACTGTTATAACAAAGACTTACAAGACAAATGCAAGTATTTGTA	64983
2008	QY	GTGACAGACTCTAGCGTTTTATCGCGTCTAGAGATAACATTAACCAAAATTGACCGAA	2067
64984	Db	GTGCAAGATTTCCAGTGGATTTATAGCTGTTTAGAAAAAATGTCAAAATATGTAGNA	65043
2068	QY	ACCAATGAGACACGGGAGCGTTTTTCCAAAGTGTGTGCGGGGAGCTATGGGTCAGCTCA	2127
65044	Db	ACTTTAGATGATGAAAAAGCTTTTCACTTTGTAGTCTGAGAGACTATGGAATCAGCTCT	65103
2128	QY	AAATTTAGCCATGACCATCTGTAAGGACACAGGAAATGCTGTGCGAGCGAGTAGCCGTAGTA	2187
65104	Db	AAACTTCTCTATGACTATTGTGTAAGGCCAAGGCACATCTCTTGTATAAGTTGCTATGCT	65163

RESULT 14  
SHE410493  
LOCUS  
DEFINITION  
Saimiriine herpesvirus 2 complete L-DNA sequence, strain C488.  
ACCESSION  
AJ410493  
VERSION  
AJ410493.1  
KEYWORDS  
GI:30348501  
complement control protein homolog; dhfr gene; dihydrofolate reductase; dutpase; gB gene; gB protein; gH gene; gH protein; helicase; IEL4-C gene; immediate early 14-C protein; IMP protein; major DNA-binding protein; mcp gene; mdbp gene; ORF23; ORF24; pol gene; polymerase; primase; protein-kinase; RRLarge; Rsmall; saimiri transformation-associated protein; stpC gene; terminase; thymidine kinase; thymidylate synthase; tip gene; TK gene; transcriptional regulator; transport protein; ts gene; tyrosine kinase-interacting protein; udg gene; uracil-DNA-glycosylase; v-cyclin gene; vC1-2 gene; vCD59 gene; vFGARAT gene; vFLIP gene; vGPCR gene; vIL-13 gene; viral Bcl-2; viral CD59 antigen; viral cyclin; viral FLICE inhibitory protein; viral G protein-coupled receptor; viral interleukin-13; viral phosphoribosylformylglycineamide amidotransferase.  
SOURCE  
ORGANISM  
Saimiriine herpesvirus 2  
Viruses; deDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.  
REFERENCE  
1 (bases 1 to 113027)  
AUTHORS  
Ensser A., Thureau M., Wittmann S. and Fickenscher H.  
TITLE  
The genome of herpesvirus saimiri C488 which is capable of transforming human T cells  
JOURNAL  
Virology 314 (2), 471-487 (2003)  
MEDLINE  
22918177  
PUBMED  
14554077  
REFERENCE  
2 (bases 1 to 113027)  
AUTHORS  
Fickenscher H., Thureau M., Wittmann S. and Ensser A.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (26-Apr-2001) Institut fuer Klinische und Molekulare Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten 4, Erlangen D-91054, Germany  
FEATURES  
source  
1. 113027  
/organism="Saimiriine herpesvirus 2"  
/mol\_type="genomic DNA"  
/strain="C488"  
/db\_xref="taxon:10381"  
/note="subtype C, L-DNA"  
complement (695. 1465)  
/gene="tip"  
complement (695. 1465)  
/gene="tip"  
/codon\_start=1  
/product="tyrosine kinase-interacting protein"  
/protein\_id="CAC84294.1"  
/db\_xref="GI:30348502"  
/translation="MANEGEIELTEPPEKERKDEKLSGSEETNTSSGSDH  
VPIEVNVIQSSRTDELONKSEIELFGQKLSGSEETAPSSSYSSQKQAFV  
IEGNDSTSTPQNVLTNLSLYTFEDARAGKGMVHKSEDLQSFLKYPDFR  
KPKRDLATDPGNETPPLPRANLGERQASTVRLHVESCKQPREKANERNIVK  
DLKLENKINVICLVVILAVLLVLSILHIGMS"  
complement (1783. 2091)  
/gene="stpC"  
complement (1783. 2091)  
/gene="stpC"  
/codon\_start=1  
/product="saimiri transformation-associated protein"  
/protein\_id="CAC84295.1"  
/db\_xref="GI:30348503"  
/translation="MASEPNLRYPPIETGRGPPGPPGPPGPPGPPGPPGPPGPP  
GPPGPPGPPGPPGPPGPPGPPGLFVTLNLLGIIVLLLIIVALLVLSKLVN  
"  
complement (4703. 5344)  
/gene="dhfr"  
complement (4703. 5344)  
/gene="dhfr"  
/codon\_start=1  
/product="dihydrofolate reductase"  
/protein\_id="CAC84296.1"  
/db\_xref="GI:30348504"  
/translation="MVLNLCIVADQNMIGKKGHLWPILLINDFKYFORMTTSSVK  
PELVNIMGKNTWFSIPEKRNPLRNILVSKLKEIPFHAFLAELNDKALIEQ  
KLNKLVDRVMIIGSSVYKDMYSHLKLFTVIMQSFETDTFFFIIDLNKKLLI  
EYGVPSNTOBEKGIKYKFEVYEKIVNTFLPKSHAGLTCSVKPKESYDFELS"  
6602. 10342  
/gene="vFGARAT"  
6602. 10342  
/gene="vFGARAT"  
/note="ORF03"  
/codon\_start=1  
/product="viral phosphoribosylformylglycineamide  
amidotransferase"  
/protein\_id="CAC84297.1"  
/db\_xref="GI:30348505"  
/db\_xref="GOA:Q80BR1"  
/db\_xref="SPTREMBL:Q80BR1"  
/translation="MLVHFDVSVEQLPQERLATSVFLENGPLQVTSNHTTTGGY  
VYNSPNTSLEPVLKFLSVOKLPSRVSEQLTSFSYGHPLQGSQTFPSKLDVY  
VKNSMFTPOGPENAFVERIEYRTTHIICDLKPKRHRIHQLMCLNPYGLTTF  
VIPQKLEMDVLDAESVVASPLSDPLMTWSNGLNLLGYSOYSVFRGDPAPVFW  
KNSRLA:INIDYWPSSNMLYDGLHAIEKGLIYFSSTLPSLTSVNMAHWY  
VMPFSAOIGHVIKLPSTHELSLTSFPPMKLOHNLKLCSSMPNIQSPFLVPKD  
LSHLPQKIYNGFISTALCLQTNIADETLQKTYLQIGNFTTHKQVLTQESG  
KEIGLSALAFLLDCVKTSGTIGTSTSLPVSAMKALAAVQCGARLFTVAGC  
POVYKLAPFSNNKAKKQHPFNLLTSDLI:IAIKNNVIDAALACTACVWAG  
FKKIGILLVNLGTEVVDNATERYPIILRDFRFPKPSLPSPINIVSEKNIWQEL  
NLRDITIKLEHPSVGCKEFTVNHNTDKLVSKVARTSIVGPMQIPVSDYSILVPMHC  
TTDEVENPDYETDIDHCVFEBATVAGTSCSAIGESTILTQADLRVTAIRATEIL  
NLSMPWNNTGNIQISLTPYTAHASTSGLVMETAKTFCALRVSCFTTANTTEG  
CASIVASAIVNTLDVSKITPDLKFNSFLLTTEKDYSLFGSAQOILEKTFIGE  
PSATSPVMLKMLSVLQTLKNDENVSHVDVSGGLIATLAEMALSGKGRVYVPHG  
RDAELKCSFTPGVIEVPGSKMYVQOFLHSENIQIIGESTSSPTLSQNLTKL  
VHEPLFKSARAFSDACPCQJHPRVMOJATVPKCPVCPGVHIVITVLLPN  
SVPLGLNAREAGQPLUSIHQPSKTYVDPHTVWGFVIGASVQDSQVCMRAL  
IAQKSHVAFORDURTMLAKPDFSVAIGALACELLFNKVIKNSQDSQVTCVNS  
SGKFSRWSNIYIPESKTAIAFOSKLSNLPCTOGTHLKFYHFKPMLEKRESGMS  
SMFYCHLSLSPQAQNYPLTPNGENAIAGVCSADGRHLALHDPDLRSLF"  
/join(10636. 11497, 11691. 11737)  
/gene="CCPHs"  
/join(10636. 11497, 11691. 11737)  
/gene="CCPHs"  
/note="alternatively spliced"  
/codon\_start=1  
/evidence=not experimental  
/product="complement control protein homolog"  
/protein\_id="CAC84298.1"  
/db\_xref="GI:30348506"  
/db\_xref="GOA:Q80BR0"  
/db\_xref="SPTREMBL:Q80BR0"  
/translation="MPSVHYICIAFLCAICFAWTLNCPKRYDSRVYVNLTSRLPY  
PTGILMTCKRGYTKGLVQSMVCSWTPQNCNKRCSPTKELLHGRYDIPODLY  
GNSMTYIKGVMILIGNTSTCLINEGQMOWIPKPLCEIRKCAAPKILNGHSH  
VKDYSYLDYSVKYCNCKIKTLIGPSKYCSETGYWPFPEETCEFTFCBKPKVANG  
NIRGNSAMYLHSQFVTSQNSGYRLAQGTENVCKNGWSPAIPTCEEVAPRGDMPH  
INTSEDPYTSSECAECPCGSYTISS"  
10636. 11718  
/gene="CCPH"  
10636. 11718  
/gene="CCPH"  
/codon\_start=1  
/product="complement control protein homolog"  
/protein\_id="CAC84299.1"  
/db\_xref="GI:30348507"  
/db\_xref="GOA:Q80BQ9"  
/db\_xref="SPTREMBL:Q80BQ9"  
/translation="MPSVHYICIAFLCAICFAWTLNCPKRYDSRVYVNLTSRLPY





Qy	1294	TCACCTGAACGTTTCGACCCAGGTTTACTCTTTATCAACAAGTTTGTGTAAACACAGCCAGTG	1353
Db	63973	ACTACAGACACATCAACAAAAGTAAACATATCTTACAAAATATGTGTAAACACACATATATT	64032
Qy	1354	TCCTTTAAATCGAAAAACAAAAATGTCATATGTGGTTTTCAGGGTACGTACTTTTGAAGTTC	1413
Db	64033	TCCTCTCAACGGGAAACTTAAAGGTGTGTGCGGTTATGTAGGCATTTATAAAAATTTT	64092
Qy	1414	AAAAGAATCCTAGACAGCGAACTCTTCGTGGAGACGCAATTCGCAAGACCGTCCGAATAC	1473
Db	64093	AAAAAATACTGGAAGTGAATCTTTTATAGATAGTCATGTCTAAGCGACCAACCTGGAATTT	64152
Qy	1474	GTATACGGTTCTTAAACAGCTACGTACACGCGCATGTACTCGTTTTCACGCGTACGGC	1533
Db	64153	GTTTACAGTTTTTTGTGTACAATACCTTTATAATCTTTTATATAATTTTCAATATAGGC	64212
Qy	1534	GTGACTAGGTTCACACAGAGAAATATCTGCAAGACCTAAAAATTTGCAACCCCTCCCGCCGCT	1593
Db	64213	GTGACTGAGAAAAATGAGTCTTAACTTAATGACTTTAGCAAAATTTAAAGTTACCAGAGAA	64272
Qy	1594	CTGGCAACCGGCGTG-----TAGACCTTCAACCGTTCGTGAGAGCTTAAACCTGGAA	1647
Db	64273	CTAACCCATTTGTATACACAAACAGATTTAGATATTTGAAAGAGAGACACTCATGTTAGAA	64332
Qy	1648	GACGACATCTTCTACACAGTGTGACGTCCCGCGCCACCGCGGGTATCACCTCCCTCCAG	1707
Db	64333	GACGATGTGTTTATCATATGTTTCTCTCCACCAACTGCCTCATAGCTTCTTGCT	64392
Qy	1708	GTTTTGTGCACACGTACTCGGCCCTTAAAGAGAGTGTTCGCCCTCAGAAATTAAGGTGGCG	1767
Db	64393	TGCTTATAAGCTGGTATCTATGCACCTGAAAGACATATTTATATCTTCGTTTAAATTAGCA	64452
Qy	1768	TGTCGCTGTTTTCGCGGGAGTTTGAGAGGAAACGTTTTTCGCGCTTTACGGTTAAACATG	1827
Db	64453	ACAACTTGGTTTAGTATAAATTTTATAGATAGAGAGTTTCAATCATTTCACTATTAAATG	64512
Qy	1828	GTGTAAGGACGGAGTTGACTTTGTCTCCCTTCGAACGTTCTCAACGGGCTGTGGCG	1887
Db	64513	CTAGTCAGAGATAATAAGAAATTTACATCTATAATGGAAGACTACATGGACTTTTAGAA	64572
Qy	1888	TTTGTCATCGACGTTTGAATCGTATATAAATTAAGGGGTACACGTTTTTACCGTAGCGTTC	1947
Db	64573	TATGCTTCTACAGTTGAATCATACAAATTAAGGCTTACTTTTTTACAGTCAACTTT	64632
Qy	1948	GGTCGCTGTACAGGGTTTGCCACTCAGTGTAGCTCAGGAAGAGATGCCCTCCCTGGTC	2007
Db	64633	GGAAGACTCAACTACTGTTTATAAGCAAGACTTTACAGACAAAAATGCAAGTATTGTA	64692
Qy	2008	GTGCAGACTCTAGCGGTTTTATCGCGTCTAGAGAAATAACATAACCAATTTGACCGAA	2067
Db	64693	GTGCAAGATTTCCAGTGGATTTATAGCATGTTTAGAAAAAATTTGTCAACAAATTTGTAGAA	64752
Qy	2068	ACCATGAGACGGGAGCGTTTTTCCAAGTGTCTGTGCGGGGACTATGGGGTCAAGCTCA	2127
Db	64753	AATTTAGATGATGAAAAAGCTTTTACCTTTGTAGTGTCTGAGACATATGGCATCAGCTCT	64812
Qy	2128	AATTTAGCCATGACCATCGTAAAGGCACAGGGAATGTCTGTGGAGCGAGTAGCCGTAGTA	2187
Db	64813	AAACTTGCCATGACTATGTAAAAAGCCCAAGGCACATCTCTTGATAAAGTTTCTATATGT	64872
Qy	2188	TTTTGGTCCACAGACGTTCCAGACAGCCAGTGTATGTAGCAATATCAAGAGGCTGT	2247
Db	64873	TTTTGTAAACCATAAAAAATAAAGTAAGTCAATATATGTTGCAATTTCAAGAGCCACA	64932
Qy	2248	AACTCAAAATTTTGTGTCATGGACAGCAACCCCTTTAAACCCCTCTCTAGAGAAACAGTTC	2307
Db	64933	AAACCTAACCACTTGTATTGACTGTATATCTTTAAAACTATTATGTCAATGACACA---	64989
Qy	2308	GATTAACCTCGGCAAGCATATAGTCCGCGCCCTCCACAAACCCCAACCAACCCCTCATC	2367
Db	64990	CAGTCTATTTCTTCTCAACATATTATTAAAGGCTTTAAATTAACCCCAACACTATTATGTC	65049

QY 6111 CGGCTTTCCATGTCGTTCAAAGGTPAGTCACGATGGATCTCCAGTTAGACTCTGTCTCCCT 6170  
Db 1089 GAAACCCGAGTGTGTACCGAATGCCAATCTGTGCCCTCCAGCTCTCAAAATTTTCATC 1148  
QY 6171 GTTAAACACCTTAACTCGAATGGAACAGCCATACACACCGTGAACCTCTCTGATGTCTCT 6230  
Db 1149 TCCAAATACCGGAATTCGGATACACACTCACTGTTGATGTCATGATGACCTAGGTCTC 1208  
QY 6231 AAAAATAACCCCGAAACAGAGCTAAATACCAATGACTGTCAACCCCTACCCCAAGCACGC 6290  
Db 1209 CCCACCAACCCCATAGGACCCAGCTACAGCTTATCTCCCACTAAATACCGGACAGCTA 1268  
QY 6291 CCCCGTACTATT-----AGACCAGGCTGAGTAAACCAAGCTATCT 6330  
Db 1269 CCGCGACTCATTAAGCCCGCCGAGAAACAGTAGTGGGTGGCAATGACAGCTCCCT 1328  
QY 6331 TTTAAAAACCCATAGTGGAGTTTGAAGGTAAACAGCT-CGTGTATATTTTCAACAGCCCTG 6389  
Db 1329 TTTAAAAAGTCAACCTTACTCGCAAGGGTAGTCTGTTGTGAGAACTACTGTCCGAGCAGC 1388  
QY 6390 TCATAAAATGATACCGACGACATCAGGTAAATTAACCTTTTTTATTCAGAGTTACGGA 6449  
Db 1389 CACAAAAATGGCGAAGATGACAAAGGTAAAGATCGACCC--TTTTTATGTATACTGAAAC 1446  
QY 6450 GGGCGTGTTTACAGTGTGTAGGTGGGAGCATATCTGCCAAATGGGAATGGCTGATGAAT 6509  
Db 1447 ATCGGTGTTTACATGATGTGTAGGTGGGAGCAGAGTTGCCCAAGCTCTACGTCCGAACAGT 1506  
QY 6510 CCACATGATGCTCGGATCAGGTGCTCTGTGTACTTACTAGTGGCGATGAACCTTAAT 6569  
Db 1507 CGGCTGTACGGGCTCTTATTAAAGTGTTCGGTGTACTTGACCAAGCGCGGAAACCTTAGT 1566  
QY 6570 TTGGCGTGTTAGTCTGTACCAACCAAAAAAAGCTTCCGTGGCTCGCTTTTGAAGCTCTA 6629  
Db 1567 TGGGTCTGTACAGTCTGTACCGACCAAAAAAGGATCGGGCGTGTCTTTCAGAGAGTTA 1626  
QY 6630 TGCAGCTTTAATCGTCTCAATAGCTTGTCTTTAAAAATGCAACTCTGTGATCACTTTA 6689  
Db 1627 GGGACGTGCTGATATGTGACACAGCTTCTGCTCGTAAATGCAACCGCTGTGTAATCTGAA 1686  
QY 6690 TGACAGTGTTCRAAAAAAACAAGATTTTAAACACAGTTAACTGCGTCTCGTGTAGT 6749  
Db 1687 CGACAGTGTCCAAAAAACAAGGTTCAGCTGACGTTTAAATCTGTATCTCGAAAGT 1746  
QY 6750 CCTCGTATTAACCTCTCAACAGAAAAAATTTTTTAAACCAACTCGCCAGGTACTGAA 6809  
Db 1747 CCTCGTAAATGACAGTTTCTACCAAGAAAAAATTTTTTACCAAGTGGCCATCCACTGAA 1806  
QY 6810 ACGATGTAACGACAGTGTGAAAAGTGTCTATCATCTCTTCACTCCCTCTTGAAG 6869  
Db 1807 AGGAGGAGCACAGTCCCCTGTGTGCTGTGTAGGATATCCCTAACTTGGAGCGGAGAC 1866  
QY 6870 TTTTGGTTACACCGACCATGTGAGATAGGCACACAGCTCAGGGGGAGGTGCGATCGTGA 6929  
Db 1867 GCGCGGACGCTCCCAAAATGGGAGAGGCACCACTCTGTGCACTCGCGGTCTGGGTT 1926  
QY 6930 GGGGTAACAATTCGTGGCGTGGGGGTACTCCAGTTTTAGGATCTCCGGCAGCACCGGTA 6989  
Db 1927 CTGATTCAGGGCGCGCTGGGGGTATTGGAGGTCAAAACTCTGGGCAGTCCCTTAA 1986  
QY 6990 CGAGTCTTCTGTAATTCCTACGACGCCGATCCACCGGGCAGTACTGAGTTAATAA 7049  
Db 1987 TGAGCTCTCTCAAAACCTATGACAGCAGGTCCTAGTGGCAGCATGCCGTTAATA 2046  
QY 7050 GTCCCCGAATTCATTCGTTGTAATTTTGTGGAGCTCTTTAGACAGTAAACCATGTTC 7109  
Db 2047 CACCCCTTATCTTGTGCTGCAAGTTTGTACAACTGCTCAGGGAATAGCCAAATTCG 2106  
QY 7110 TCCGACATGGGTACGATGAACCGCTCTCTGTGCTGTGTCATGAGCACCTGGAGTA 7169  
Db 2107 CCCTAGCCGCGGAAACAGGTACGGCTCGCTTTGTGGTGTGTGACCAATATCTGAATGG 2166

QY 7170 TCTTAGACAGGTATAGGGCCGTTCCACTTAAGGTTGCCGTTGGTAGACGGATTGC 7229  
Db 2167 TCTTTGCAAGGTATAGGGTCTTCTCAAGCTTTAGAGGGGTACGTGGCAGTCTCGATTGA 2226  
QY 7230 GGGCAATTACCGTCAAGTGTCTCATTTTCCACGACTATCATCCGAGTCCGCGATAGTGA 7289  
Db 2227 GGGTGGCAGCAGACAGGTAICTAACTCTCTGAAGTACTGATCCAGGACGGTAATGAT 2286  
QY 7290 ACTTGTAATATGATTTAAAAAGATGCTTTCCCGTAAATGGCGACTCATGGTCTGTTT 7349  
Db 2287 ACCTAAAACAGATGGTTGAACAGGTGATCTTTAAGGGGCTTCTCGATGTCTATTGTA 2346  
QY 7350 A-----ACGTTTCAAGGGTGTAAAAAATCAGAAAGTCTCGCTGGGACAACTT 7398  
Db 2347 CTATGACACGCCACTCTCTCTTAGGTTAAGAGCTTCGGCGGCTCTGTGTGGAAGCTT 2406  
QY 7399 TTGCGGACTCTCCGAACCGCTCAGGATAAGCTGTATACACCGGTGCTTCTTTATAACGA 7458  
Db 2407 CGTGGCCTCTCGAGCAACTGAAGGCCAACTCTACCAAGTGTGTGCTCTTTATAAATGA 2466  
QY 7459 AGCATACCTCAAAAAACATGACCCACGTCACCTAAATAAGCTTCACAGAAAGTCTGTCG 7518  
Db 2467 CGCATCAAAACATCTACGATCCAGTGCCTTAATAGATGTTGGAAGATGTGTGAT 2526  
QY 7519 ATGTATTGTAATGGAATGTCTCCCTGGGTCCCATCTCCGGGTCTATCGGGATCTGAA 7578  
Db 2527 TCGGATTATGAAGAATGTTCCAAAGCTTGGTGGCTATGTGTCTGTTTTACAGACATTAA 2586  
QY 7579 TCTCTTTAACTGTCTGCTCTACCGGGTTCGGGGTTAAACCCCGCGGGCGCCAC 7638  
Db 2587 CATGTTTAACTTTTCTGCTTTTTCGTGCTCTCGAATGAGGACCAAGGCGCGCGG 2646  
QY 7639 CTGTAACTCCGTCGCGAGAATCGCCACAGGATCTGTAGAAATTTCTACCGAGAGGC 7698  
Db 2647 GTACAACTGTCATGCGCAGAGGATCCCAAGGATTTATTCGGATCTCTCAGGAGAGAT 2706  
QY 7699 CTTTGTGTTACCGAAGATGTTTCAAGCTCTGCTCGCAGCGGGTCTGTTATACCGCC 7758  
Db 2707 CTATTCTGCAAGAAAAGGCAATTTCTGACAGCGCATGACGCGGGTGAACCTGCTTCC 2766  
QY 7759 CCAGTTGCCAGAGTTCTGACAGAGCTCTACGCGAATAAGAGCCCAAGTGTTTGGGGC 7818  
Db 2767 AGCCATATGAAGTACTTACAGAAATATACACTGAATGAAGGCCAAATTCCTTGGGGC 2826  
QY 7819 GTGGGTCTCTATATGTTGACGGGCTCTATGATGCTATTTGCGGCTATGCGGCTCTGCT 7878  
Db 2827 CTGAGGGGACTCTGCTGCAATCGAGGCCCAATATGATATTAACTCTTCCCTACTGAA 2886  
QY 7879 CACTTATATACCTTGGATCGGAGGAAAACTTGAACCTTAGACTCAAGACACTGTGCA 7938  
Db 2887 GCTCTACACAGTACGATACGCGGGTCTCTCTGAGCAGTCCAGGSCCTCTGCT 2946  
QY 7939 ACTTGTCTTCAACCATCTTTCTTCAAGAAATCTTAGCCCTTATGACGCTACTGCGCAA 7998  
Db 2947 TTTGTTTTCACACCGGTCTACCTCCAGGATTTATGGCGCGCTGGAGATCATGACCAA 3006  
QY 7999 CGGAGAGATGTTCTCTGACAAATTTTATCCATCACCGGTACGCGCGAGAGAGCGGCC 8058  
Db 3007 GGTTCAGTCTCGCCCTGAAAACCTTTTACAGCATCACCGGTTCTGCTGAGAAACCGCGCC 3066  
QY 8059 TGCTGTACTGGAAGTACTAGCGGATGACGTGTCCGGGAAGCAGCTTGTCCCGATTC 8118  
Db 3067 AATTACCACCGCAAGGTCACTGGACTGTCTATTCAGAAAGCGCTCTCATGCCAGAAATC 3126  
QY 8119 CTTATCTGTCAGTATCGGAACCGGGCTTCTCCCGCACCCCTGGTTGACCTCAGTAA 8178  
Db 3127 TTTAATTTTGCAATCTTGGAGCCAGGACTGTTGGCGGCTTCCATGTTAGACCTCAGCGA 3186  
QY 8179 TGCTTTAGAAATCCAGAAATCATCTCAGCGGCCCCACCTCTGAGTCAATTTGTATCAC 8238  
Db 3187 TGTGCTGGCAAAACCCCGGTTATTTCTGAGCGCCCTGCTCCCTGAGCCAGTTTGTATTAG 3246  
QY 8239 AAACAGCACCCCGTCTGCTCAGTCAAGTATTATTAGCCCAACCCAGGGCGTGT 8298

Db	3247	CAAACCCATCCCAACATGCCCGCACACCGTCAGCATCATCCCTTTAAACC---CATCGGG	3303
Qy	8299	TCCCGGCAATGTTTATGTCACACGTCGGAAGCGGTGCACAGAGCATTCACACCGGC	8358
Db	3304	TACAGACCGCGCTTTATTAGTACGTGGCAGCGCGTCACAGATATGGTGTACACAC	3363
Qy	8359	ACAGAC	8364
Db	3364	ATCCAC	3369

Search completed: April 11, 2004, 04:15:49  
Job time : 22587 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 16:16:38 ; Search time 2057 Seconds  
(without alignments)  
18737.920 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038

Perfect score: 9073

Sequence: 1 atggagagttcgtcgatg.....ggggccacgcatgccttga 9073

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9073	100.0	128139	6	Aa164291 RRV genom
2	9073	100.0	133719	3	Aac64754 Macaca mu
3	1607.6	17.7	137507	2	Aav19941 KSHV long
4	1002	11.0	35100	2	Aav73803 KSHV LUR
5	714	7.9	35100	2	Aav73804 KSHV LUR
6	200.8	2.2	7361	2	Aav62153 HSV-2 str
7	200.8	2.2	117213	2	Aav62176 HSV-2 str
8	200.8	2.2	154746	6	Aad25519 Human her
9	199.2	2.2	8952	2	Aav62130 HSV-2 str
10	154	1.7	125157	5	Aah74202 Nucleotid
11	153.4	1.7	2871	2	Aat85076 Human cyt
12	153.4	1.7	2871	2	Aat85062 Human cyt
13	153.4	1.7	5290	8	Acc70010 Nucleotid
14	153.4	1.7	229354	6	Abq74179 Human cyt
15	152.4	1.7	124884	5	Aah74201 Nucleotid
16	152.4	1.7	124884	7	Ada14878 Human her
17	137.6	1.5	1532	4	Aas63241 Escherich
18	129.4	1.4	681	7	Acf71966 Phototrab
19	129.4	1.4	110000	7	Continuation (52 o
20	129.4	1.4	110000	7	Continuation (4 of
21	125.8	1.4	7316	3	Aac62115 BamHI "G"
22	121	1.3	657	7	Abz40183 N. gonorr
23	121	1.3	714	7	Abz40182 N. gonorr

C 24	117.6	1.3	657	5	Aaf94351
C 25	117.6	1.3	657	6	Abk64915
C 26	117.6	1.3	110000	2	Aat42063 Haemophil
C 27	114.6	1.3	110000	3	Continuation (12 o
C 28	114.6	1.3	349980	3	Aaf21609 Neisseria
C 29	114.6	1.3	349980	3	Aaf21610 Neisseria
C 30	111	1.2	4161	2	Aav42547 Kpni-Kpni
C 31	109.4	1.2	4161	2	Aav42517 Kpni-Kpni
C 32	107.4	1.2	1283	4	Aas09498 Atlantic
C 33	107.4	1.2	1355	4	Aas09499 Atlantic
C 34	96	1.1	4100	2	Aaz32019 Human MET
C 35	96	1.1	4100	5	Aac90076 L13855 CD
C 36	95.4	1.1	110000	6	Continuation (5 of
C 37	95.4	1.1	110000	6	Continuation (4 of
C 38	94	1.0	696	9	Add26226 Pseudomon
C 39	94	1.0	696	9	Add26228 Pseudomon
C 40	91	1.0	119211	4	Aaf28553 Genomic f
C 41	90.2	1.0	651	4	Aah52516 S. epider
C 42	90.2	1.0	657	6	Ahn93020 Staphyloc
C 43	90.2	1.0	2975	4	Aah53981 S. epider
C 44	90.2	1.0	4114	4	Aah54820 S. epider
C 45	90	1.0	1967	4	Aah34344 Human col

## ALIGNMENTS

### RESULT 1

AA164291  
ID AA164291 standard; DNA; 128139 BP.

AC AA164291;

DT 22-APR-2002 (first entry)

DE RRV genome nucleotide sequence.

KW RRV, rheus rhadinovirus; Japanese macaque virus; multiple sclerosis;  
JMV; cytostatic; antiasthmatic; antiallergic; dermatological; vulnery;  
gene therapy; leucopenia; thrombocytopaenia; inflammatory disease;  
asthma; allergy; dermatitis; virus; ds.

OS Macaca mulatta rhadinovirus 17577.

Key Location/Qualifiers

FT CDS 1353..2674

FT CDS /\*tag= a

FT CDS /product= "RRV R1"

FT CDS complement(2692..3258)

FT CDS /\*tag= b

FT CDS /product= "dihydrofolate reductase"

FT CDS /label= RRV ORF2

FT CDS /note= "has\_similarity to Kaposi's sarcoma-associated virus (KSHV) open reading frame (ORF) 2"

FT CDS 3676..5613

FT CDS /\*tag= c

FT CDS /product= "complement binding protein"

FT CDS /label= RRV ORF4

FT CDS /note= "has\_similarity to KSHV ORF4"

FT CDS 5035..9443

FT CDS /\*tag= d

FT CDS /product= "ssDNA binding protein"

FT CDS /label= RRV ORF6

FT CDS /note= "has\_similarity to KSHV ORF6"

FT CDS 9468..11528

FT CDS /\*tag= e

FT CDS /product= "transport protein"

FT CDS /label= RRV ORF7

FT CDS /note= "has\_similarity to KSHV ORF7"

FT CDS 11515..14004

FT CDS /\*tag= f

FT CDS /product= "glycoprotein B"

FT CDS /label= RRV ORF8

[illegible]

FT	/note= "has similarity to KSHV ORF41"
FT	complement(58525..59343)
FT	/*tag= an
FT	/label= RRV ORF42
FT	/note= "has similarity to KSHV ORF42"
FT	complement(59297..61027)
CDS	/*tag= ao
FT	/product= "capsid protein"
FT	/label= RRV ORF43
FT	/note= "has similarity to KSHV ORF43"
FT	
Query Match 100.0%; Score 9073; DB 6; Length 128139;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGGAGAGTTCCGTCGGATGGACCAAAACACCTCGAAACAAATCCGGGGTTCATCTTGAAC 60
Db	55386 ATGGAGAGTTCGTCGGATGGACCAAAACACCTCGAAACAAATCCGGGGTTCATCTTGAAC 55445
QY	61 ATGACGTCGGATGCCAAAGTCAGGGGTCTCGTGATACGTCAGTCGCGCTGTCAAAATATA 120
Db	55446 ATGACGTCGGATGCCAAAGTCAGGGGTCTCGTGATACGTCAGTCGCGCTGTCAAAATATA 55505
QY	121 ACTACAGCCACCGGAAATGGTGTGGTACGACCTGGCTTCGATCCGGCTGAAGACTCC 180
Db	55506 ACTACAGCCACCGGAAATGGTGTGGTACGACCTGGCTTCGATCCGGCTGAAGACTCC 55565
QY	181 GGGCGGTTCCTGCGTTTACCGTTTATCTAATTACGGAACTGCTGTGCTCGGAAAGT 240
Db	55566 GGGCGGTTCCTGCGGTTCACCGTTTATCTAATTACGGAACTGCTGTGCTCGGAAAGT 55625
QY	241 ACCAGCATATCGGCCCTGTACCAAATTTAAACTGCCTGATCAGGGGCGCACCAACATA 300
Db	55626 ACCAGCATATCGGCCCTGTACCAAATTTAAACTGCCTGATCAGGGGCGCACCAACATA 55685
QY	301 GCCGCACAGAACTTATCGCGTCGCCCTAAAGAGTTCTGTCCACGATTTT CAGGCGTTTT 360
Db	55686 GCCGCACAGAACTTATCGCGTCGCCCTAAAGAGTTCTGTCCACGATTTT CAGGCGTTTT 55745
QY	361 GGGTTTAGAGCGGCACATCAATATAGCCGTCAGAAAGCTCATCAGACGGAGCGGTA 420
Db	55746 GGGTTTAGAGCGGCACATCAATATAGCCGTCAGAAAGCTCATCAGACGGAGCGGTA 55805
QY	421 TCCATAGACAAATTTCAGCAACAGAGAGCTATCGAAGTATGCGCCGGTTATAGTGACATT 480
Db	55806 TCCATAGACAAATTTCAGCAACAGAGAGCTATCGAAGTATGCGCCGGTTATAGTGACATT 55865
QY	481 ATGAAGAGGTTATGCGCAAAAACCCAAATGCGATGTACGGAGCTATATCCAACCGGAAT 540
Db	55866 ATGAAGAGGTTATGCGCAAAAACCCAAATGCGATGTACGGAGCTATATCCAACCGGAAT 55925
QY	541 TTTGAAACCTCTCGAGATGACCGACCGTGTATTATGACATTCCAATATTATTGTAATC 600
Db	55926 TTTGAAACCTCTCGAGATGACCGACCGTGTATTATGACATTCCAATATTATTGTAATC 55985
QY	601 GACGAGCGCGAACCCCTGCTCTTTATCATACTCAACACCGTGTGTTTCTTTTACTGGTTC 660
Db	55986 GACGAGCGCGAACCCCTGCTCTTTATCATACTCAACACCGTGTGTTTCTTTTACTGGTTC 56045
QY	661 CTAAACAGCTGGCTAAATACCCCTCTTTACGCCAGGGGGCGGTTCCGTCATAGTAGTC 720
Db	56046 CTAAACAGCTGGCTAAATACCCCTCTTTACGCCAGGGGGCGGTTCCGTCATAGTAGTC 56105
QY	721 GTCGGTTTCGCCAACGGCAGACAAACCGGTTCCAGTCAACTTACAAACACGGGACGCAAAAG 780
Db	56106 GTCCGTTTCGCCAACGGCAGACAAACCGGTTCCAGTCAACTTACAAACACGGGACGCAAAAG 56165
QY	781 ACGGAGATATGTCGTGCGAGAACATCCTTAACTTATGATAGGGAAGAGTCTGTAICT 840
Db	56166 ACGGAGATATGTCGTGCGAGAACATCCTTAACTTATGATAGGGAAGAGTCTGTAICT 56225
QY	841 GAGTACGTACACTTGGAGAGAAACTGGGCGCTGTTTATAAACAAACGAGCGCTGCACCGAT 900



1981 CTTGAGGAGAGATGCCCTCCCTGGTGGTGGAGGACTCTAGCGGTTTTATCGGTGCTTA 2040  
57366 CTTGAGGAGAGATGCCCTCCCTGGTGGTGGAGGACTCTAGCGGTTTTATCGGTGCTTA 57425  
2041 GAGAAATAACATAACCAAAATGACCGAAACCAATGAGGACGGAGCGTTTCCAAAGTGTGC 2100  
57426 GAGAAATAACATAACCAAAATGACCGAAACCAATGAGGACGGAGCGTTTCCAAAGTGTGC 57485  
2101 TGTGGGGGGAGCTATGGGGTCAAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 2160  
57486 TGTGGGGGGAGCTATGGGGTCAAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 57545  
2161 ATGTGCTGGAGCGAGTACCGTATGATTTGGGTCCCAAGAAACGCTCCAGACAAAGCCAC 2220  
57546 ATGTGCTGGAGCGAGTACCGTATGATTTGGGTCCCAAGAAACGCTCCAGACAAAGCCAC 57605  
2221 GTGTATGTAGCAATATCAAGGGCTGTAACTCAAAATTTTGGTCAATGACGACACACCCC 2280  
57606 GTGTATGTAGCAATATCAAGGGCTGTAACTCAAAATTTTGGTCAATGACGACACACCCC 57665  
2281 CTTAAACCCCTCTCAGAGAACCGTCTGATTAACACCTCCGCAAGCATATAGTCCGGGCC 2340  
57666 CTTAAACCCCTCTCAGAGAACCGTCTGATTAACACCTCCGCAAGCATATAGTCCGGGCC 57725  
2341 CTCACAAACCCCAACCAACCTCTACTACTAAATAAATAAACAAGCAAAATGGAAC 2400  
57726 CTCACAAACCCCAACCAACCTCTACTACTAAATAAATAAACAAGCAAAATGGAAC 57785  
2401 ATGTGTTTTTATCAGTCCCAACACCGCCACGATAGTGTCTATTTCCACACACCGGG 2460  
57786 ATGTGTTTTTATCAGTCCCAACACCGCCACGATAGTGTCTATTTCCACACACCGGG 57845  
2461 GGGTGGCGCCATAGTTGAGCAACAGAAACCGCGGCCCTCTGGGTGAGAGTAGTCC 2520  
57846 GGGTGGCGCCATAGTTGAGCAACAGAAACCGCGGCCCTCTGGGTGAGAGTAGTCC 57905  
2521 CGCCCCCGGTGTGAGTGCATACGTCTCTTAGGGCCCTGGTGTATACCAACATAAATCA 2580  
57906 CGCCCCCGGTGTGAGTGCATACGTCTCTTAGGGCCCTGGTGTATACCAACATAAATCA 57965  
2581 CGTGTGTATCAGCTTGCAGTCCCTCGGTGAGCGCGCCCTCTGCGAGCGGACGACAC 2640  
57966 CGTGTGTATCAGCTTGCAGTCCCTCGGTGAGCGCGCCCTCTGCGAGCGGACGACAC 58025  
2641 CGCGCTGTGTAACGATCTGCATGCTGCTGCTGTCTGAAGATGACACAGTCAATCA 2700  
58026 CGCGCTGTGTAACGATCTGCATGCTGCTGCTGTCTGAAGATGACACAGTCAATCA 58085  
2701 CATCAGTGGCGCGCGCTCATCGGCTGCTGATTCGCTCCATGCTCAACAGCGCTC 2760  
58086 CATCAGTGGCGCGCGCTCATCGGCTGCTGATTCGCTCCATGCTCAACAGCGCTC 58145  
2761 GGTGTCCACTTGTGCTGTCTGAAATAAATGCTCTCTGTCTGCTGCAAGTTGATCCAG 2820  
58146 GGTGTCCACTTGTGCTGTCTGAAATAAATGCTCTCTGTCTGCTGCAAGTTGATCCAG 58205  
2821 CAACAAAGGTTCTTCTCATCGGAGACCATCGCCCTCTGCGGTGATCTAGATATAT 2880  
58206 CAACAAAGGTTCTTCTCATCGGAGACCATCGCCCTCTGCGGTGATCTAGATATAT 58265  
2881 CACTGACACGCGCGGAGCCCGTGGCGAAAGCCGTCAGCTTCTCCGCCACTTTCAT 2940  
58266 CACTGACACGCGCGGAGCCCGTGGCGAAAGCCGTCAGCTTCTCCGCCACTTTCAT 58325  
2941 CAGTGTCTGTGTGATCTATAAATAAATGCTCTGCTGCTGTGATTCAGATGATGGG 3000  
58326 CAGTGTCTGTGTGATCTATAAATAAATGCTCTGCTGCTGTGATTCAGATGATGGG 58385  
3001 GTGCGCGGTGTGGCCAGTAAACCAACCGGCGCGGACCAACCAACCGCGGTGATG 3060  
58386 GTGCGCGGTGTGGCCAGTAAACCAACCGGCGCGGACCAACCAACCGCGGTGATG 58445

3061 GCGCGGGCGCGCTCGGTGGTGGTGGCACTGTGCGAAGTAATGACGCTAGGTAGCGGC 3120  
58446 GCGCGGGCGCGCTCGGTGGTGGTGGCACTGTGCGAAGTAATGACGCTAGGTAGCGGC 58505  
3121 CGCTCTGTTTAAAGGACTCACCGTCCCGTGGTGGGATCTATTTCTAGGTGGGAGAGAGA 3180  
58506 CGCTCTGTTTAAAGGACTCACCGTCCCGTGGTGGGATCTATTTCTAGGTGGGAGAGAGA 58565  
3181 CGCGCGGTTCCTCTCTTCATCCACATCATGTAGTGGTCTCCGAATGAGTCCG 3240  
58566 CGCGCGGTTCCTCTCTTCATCCACATCATGTAGTGGTCTCCGAATGAGTCCG 58625  
3241 TTTGCTCTGGCGTTTTTGGAGTAGTGGGAGGGAATCTTGGTCAAAAAACATCTCTCAGGTG 3300  
58626 TTTGCTCTGGCGTTTTTGGAGTAGTGGGAGGGAATCTTGGTCAAAAAACATCTCTCAGGTG 58685  
3301 CAGTCATGTCAATCACTGGGGGGGCTCGGTACCCGGGCGTACGGGCAACCCGCTACGGG 3360  
58686 CAGTCATGTCAATCACTGGGGGGGCTCGGTACCCGGGCGTACGGGCAACCCGCTACGGG 58745  
3361 ACCCGGGGAAACAAAAACGCTCTTGTCTCTTTTCTAGTGGTCCCGGGAATCGGACGCA 3420  
58746 ACCCGGGGAAACAAAAACGCTCTTGTCTCTTTTCTAGTGGTCCCGGGAATCGGACGCA 58805  
3421 TCCTGGAGTGGCGGGGCTCTTGAACGGTGGATAGACATAGCATGCGCGAA 3480  
58806 TCCTGGAGTGGCGGGGCTCTTGAACGGTGGATAGACATAGCATGCGCGAA 58865  
3481 CGTTCACCTGTAAACGCACTGCTATFCCCGGATCAGTTAAATTTCCAGTCAACGCGACC 3540  
58866 CGTTCACCTGTAAACGCACTGCTATFCCCGGATCAGTTAAATTTCCAGTCAACGCGACC 58925  
3541 CCTCGGATCTGAACTAGATAGTCAATTCGCAAGTTTAAATGCTTCAGCTAGAACTG 3600  
58926 CCTCGGATCTGAACTAGATAGTCAATTCGCAAGTTTAAATGCTTCAGCTAGAACTG 58985  
3601 CGGACAGGAAACCGGTTGCGGTGGGAGTGGAGGAGACGAGTGGCGGCTT 3660  
58986 CGGACAGGAAACCGGTTGCGGTGGGAGTGGAGGAGACGAGTGGCGGCTT 59045  
3661 GAGACAGATGCTCTGCTCGCTTATCAGCTGCTGCTGCTATAGCTTCTTCCCA 3720  
59046 GAGACAGATGCTCTGCTCGCTTATCAGCTGCTGCTGCTATAGCTTCTTCCCA 59105  
3721 CAGCATAAAAACGCACTCTGAAGCTTGGCAGAAAGGACGCTTATTTATGTAGTGTGAA 3780  
59106 CAGCATAAAAACGCACTCTGAAGCTTGGCAGAAAGGACGCTTATTTATGTAGTGTGAA 59165  
3781 CCAATCCAGCAAGTGGGCTGCTGACCCCGCTTCCCTTCCACCGTCAAGATGCT 3840  
59166 CCAATCCAGCAAGTGGGCTGCTGACCCCGCTTCCCTTCCACCGTCAAGATGCT 59225  
3841 GTTTAAAGCAGAACTCCCGTTTGGCCAAAGTCTAAGCAACCGTGAAGAGGAGCGT 3900  
59226 GTTTAAAGCAGAACTCCCGTTTGGCCAAAGTCTAAGCAACCGTGAAGAGGAGCGT 59285  
3901 GAACCCAGGTACAGTATGGCTATCTTTTAAATAATTTTGGAGCTTGGAGTATCT 3960  
59286 GAACCCAGGTACAGTATGGCTATCTTTTAAATAATTTTGGAGCTTGGAGTATCT 59345  
3961 GTAGTCCGAGCAACGCTAAAGCTAGTCCGTTGGTGAACCGGATGTAAGGATCTCTG 4020  
59346 GTAGTCCGAGCAACGCTAAAGCTAGTCCGTTGGTGAACCGGATGTAAGGATCTCTG 59405  
4021 GCCAAGGATCAGCACTTTAATATCTCTCGGTTCCGCAACAAATAGCAACCAATATCT 4080  
59406 GCCAAGGATCAGCACTTTAATATCTCTCGGTTCCGCAACAAATAGCAACCAATATCT 59465  
4081 ATCTATCGGTGATTAATCAACGTTTACGTTGCTCATATCAATAACCGCTTCAGCAGAC 4140  
59466 ATCTATCGGTGATTAATCAACGTTTACGTTGCTCATATCAATAACCGCTTCAGCAGAC 59525  
4141 CGCGAGCTTCTGTTTTTAAATAATGAGAGAGATTTTAAATAATTTCAAGCCATGAGTCCGTAAG 4200

59526	Db		CGGACGCTTCTGTTTAAATAATGGAGAAAGATTTAAAAATTCAGGCCATGAGTCGCTAAG	59585
4201	Qy		TAGCAACGTTTGTCGGAGGGTCTCTAAACACTTCTGGTGACATTTTACTCCAAACTAT	4260
59586	Db		TAGCAACGTTTGTCGGAGGGTCTCTAAACACTTCTGGTGACATTTTACTCCAAACTAT	59645
4261	Qy		TGTTTTAAGCCAAACCTCCATAAATCCCGCGCTCTGGTAAATTTACCTTTACGTATACGTGAA	4320
59646	Db		TGTTTTAAGCCAAACCTCCATAAATCCCGCGCTCTGGTAAATTTACCTTTACGTATACGTGAA	59705
4321	Qy		ACTTTTGTAATCTCTGCTCAGAAATCCCAACAGAGTGGCGCGTATAAGCAGTTCCCGCTT	4380
59706	Db		ACTTTTGTAATCTCTGCTCAGAAATCCCAACAGAGTGGCGCGTATAAGCAGTTCCCGCTT	59765
4381	Qy		TCACTACCGTTTATGTGAAGTCTTCAAACTCGGCTAAAACTCATCAAGACTATATTGC	4440
59766	Db		TCACTACCGTTTATGTGAAGTCTTCAAACTCGGCTAAAACTCATCAAGACTATATTGC	59825
4441	Qy		ATATAATTCMAAATTTGAAGCTCTCTCTGGGGACACTGTAAATTTAACTTTGTACAAATC	4500
59826	Db		ATATAATTCMAAATTTGAAGCTCTCTCTGGGGACACTGTAAATTTAACTTTGTACAAATC	59885
4501	Qy		CTTATAAAAACGCCCTAAGGTGAATCCATTTACACATATTTCTGTGGTAGAGCATCA	4560
59886	Db		CTTATAAAAACGCCCTAAGGTGAATCCATTTACACATATTTCTGTGGTAGAGCATCA	59945
4561	Qy		GCCTTATATCGCAATGTTGCTGTACACAGCCTTGGACATAAAGCTAAATTTGTAAATGA	4620
59946	Db		GCCTTATATCGCAATGTTGCTGTACACAGCCTTGGACATAAAGCTAAATTTGTAAATGA	60005
4621	Qy		AAAGAGCAATCTCAAAAGCGCTCCGGAAATATAAAGTCACTAAAGTGGTATTATTTACAA	4680
60006	Db		AAAGAGCAATCTCAAAAGCGCTCCGGAAATATAAAGTCACTAAAGTGGTATTATTTACAA	60065
4681	Qy		CATGGATTTACTCAAGCCGGTACAACTGTAAATTTAAAAAAAATCCACATGTATATAAT	4740
60066	Db		CATGGATTTACTCAAGCCGGTACAACTGTAAATTTAAAAAAAATCCACATGTATATAAT	60125
4741	Qy		AACTTTAAGGTATACATATCTCTCATGTTTTATACATATATGAAGCTTAATGCTGTACT	4800
60126	Db		AACTTTAAGGTATACATATCTCTCATGTTTTATACATATATGAAGCTTAATGCTGTACT	60185
4801	Qy		TATATAATATGTATAGAAATTTATGCCAATACAGTCAATGTGTAACTATTTTAAAGGCATAAT	4860
60186	Db		TATATAATATGTATAGAAATTTATGCCAATACAGTCAATGTGTAACTATTTTAAAGGCATAAT	60245
4861	Qy		GCAACATCAAAATATTAACACAAAATGCGCTTCGGATATCCGCCACAGCTGTTAAC	4920
60246	Db		GCAACATCAAAATATTAACACAAAATGCGCTTCGGATATCCGCCACAGCTGTTAAC	60305
4921	Qy		CCGAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTATTTCTCAACATATG	4980
60306	Db		CCGAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTATTTCTCAACATATG	60365
4981	Qy		CAACAAATATTTAGGGTAAACCATGTTAAATAAATACGACATGCAATTTTATAGCTA	5040
60366	Db		CAACAAATATTTAGGGTAAACCATGTTAAATAAATACGACATGCAATTTTATAGCTA	60425
5041	Qy		AGCTCTCAAGACAATTTCTCTTCAATTTGATGACTCATCAACCTCGGTTCCCGTGTTA	5100
60426	Db		AGCTCTCAAGACAATTTCTCTTCAATTTGATGACTCATCAACCTCGGTTCCCGTGTTA	60485
5101	Qy		ATGTCGGAATATATAGTCCAGGATACCTCTGTGATCCGCCACAAATTAATTTCCAGGCT	5160
60486	Db		ATGTCGGAATATATAGTCCAGGATACCTCTGTGATCCGCCACAAATTAATTTCCAGGCT	60545
5161	Qy		TGTAAATATATCAATCCAGTGAATCAGCTGCTTAATGTCGAAGATGTTGTTCTTAATTTTACG	5220
60546	Db		TGTAAATATATCAATCCAGTGAATCAGCTGCTTAATGTCGAAGATGTTGTTCTTAATTTTACG	60605
5221	Qy		CCATCCAAATGGTGAAGTGGGTGACATTTTGGGACAAAGCCCGTAAAAATTGGCAGCTCC	5280

60606	CCATCCAAATGGTGAGTGGTGACACTTTTGGGACAAAGCCCCGGTAAAAATTGGCAGCGTCC	60666
5281	ACAAACGGAAGGATGCAATTTGATCGTTTACGAGGGGCAGCATATTTTATTTTATAACA	5340
60666	ACAAACGGAAGGATGCAATTTGATCGTTTACGAGGGGCAGCATATTTTATTTTATAACA	60725
5341	TCAGGAACCGGACCGGTCAATCAGCAGCATATGTCCATTTAGATTTGGAGTCGTGACACTG	5400
60726	TCAGGAACCGGACCGGTCAATCAGCAGCATATGTCCATTTAGATTTGGAGTCGTGACACTG	60785
5401	GTTTGAATTGACGCGTGAAGAAAAATGTTTTCGTGTCCATTCGCTAACTATTTTTTAAACCAC	5460
60786	GTTTGAATTGACGCGTGAAGAAAAATGTTTTCGTGTCCATTCGCTAACTATTTTTTAAACCAC	60845
5461	ACAGAGGCGGATATAAACCGGTCCCGGGCTTAGGCCGTGAATGCGCAGCGCAGCGCTC	5520
60846	ACGAGGCGGGAATAAACCGGTCCCGGGCTTAGGCCGTGAATGCGCAGCGCAGCGCTC	60905
5521	ATTGCCATTTCAGATTTCTTAAATCTGTTCGAAATACGAGTCTAAATCAAGATATAAC	5580
60906	ATTGCCATTTCAGATTTCTTAAATCTGTTCGAAATACGAGTCTAAATCAAGATATAAC	60965
5581	CACGACAAGTTCAACAGAGAGCAACGGCGGTGGTTTCCACCGGACATCCCGCGTCCACA	5640
60966	CACGACAAGTTCAACAGAGAGCAACGGCGGTGGTTTCCACCGGACATCCCGCGTCCACA	61025
5641	AACCCGAACTCTCTAGGAACGACATCAAGAGGTTTCAGACAAAAATTCGGAACCGAAACC	5700
61026	AACCCGAACTCTCTAGGAACGACATCAAGAGGTTTCAGACAAAAATTCGGAACCGAAACC	61085
5701	CTAGTTCTGAGCACCATCTGGTTTAAACTTCGGTTTAAGAGACGAGCGGCACACATGTC	5760
61086	CTAGTTCTGAGCACCATCTGGTTTAAACTTCGGTTTAAGAGACGAGCGGCACACATGTC	61145
5761	TGCTTGTGAGGATTGAGCTTAGGTGGCTGCAGGCTTGCTCAGAGCCCGGTAGTGGCTG	5820
61146	TGCTTGTGAGGATTGAGCTTAGGTGGCTGCAGGCTTGCTCAGAGCCCGGTAGTGGCTG	61205
5821	CGCGCGCTTCGCGCACCTTTGACAGCGCACCGCAGTTCACAGCGAGGACGAGCGCTCT	5880
61206	CGCGCGCTTCGCGCACCTTTGACAGCGCACCGCAGTTCACAGCGAGGACGAGCGCTCT	61265
5881	TTGAGTGTGGGCCATATCTCGGCCATAGGAGATTCGCTCGCGTGATCTTTTAGTGGCATA	5940
61266	TTGAGTGTGGGCCATATCTCGGCCATAGGAGATTCGCTCGCGTGATCTTTTAGTGGCATA	61325
5941	TATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGAAATAAGTCCGGTCTTTTCT	6000
61326	TATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGAAATAAGTCCGGTCTTTTCT	61385
6001	ATCGTTTTCTGAAATCTATAGTTTACATCTTTTAGTATAACAAAGACGCTTAAAAACCCA	6060
61386	ATCGTTTTCTGAAATCTATAGTTTACATCTTTTAGTATAACAAAGACGCTTAAAAACCCA	61445
6061	GGATGTCGACACCTTTTAAAAATCGAAAAGCGACCGGATGGCTCTATCCGCGTTTCCA	6120
61446	GGATGTCGACACCTTTTAAAAATCGAAAAGCGACCGGATGGCTCTATCCGCGTTTCCA	61505
6121	TGTCGTTTCAAAGGTAGTCACGATGGATCTCCAGTTAGACATCTGTCTCCCTGTTTAACACC	6180
61506	TGTCGTTTCAAAGGTAGTCACGATGGATCTCCAGTTAGACATCTGTCTCCCTGTTTAACACC	61565
6181	TTAACTTGGATGGAACACGCCATGACCGTGAACTTCTGTGATGTCTCTTAAACTTAACC	6240
61566	TTAACTTGGATGGAACACGCCATGACCGTGAACTTCTGTGATGTCTCTTAAACTTAACC	61625
6241	CCGAAACAGAGCTTAAATACCAATGATGTACCCCTTACCCCAAGCCACGCCCCCGTACTA	6300
61626	CCGAAACAGAGCTTAAATACCAATGATGTACCCCTTACCCCAAGCCACGCCCCCGTACTA	61685
6301	TTAGACCAAGGTCAGTAAGCAACGCTATCTCTTTAAAAACCCATACGTGAGTTTGTAAAGT	6360
61686	TTAGACCAAGGTCAGTAAGCAACGCTATCTCTTTAAAAACCCATACGTGAGTTTGTAAAGT	61745

Qy	6361	AAACAGCTCGTGTATATTTT	CAGACGCTGTGCATATAAATGATACCGACAAATCAGGTA	6420	7441	GGTGCTCTTTATTAACGAGCATACGTCAAAAACATGACCCAGTGCATTAATAAGCT	7500	
Db	61746	AAACAGCTCGTGTATATTTT	CAGACGCTGTGCATATAAATGATACCGACAAATCAGGTA	61805	62826	GGTGCTCTTTATTAACGAGCATACGTCAAAAACATGACCCAGTGCATTAATAAGCT	62885	
Qy	6421	ATTAAACTTTTATTAACAGATTCAGAGGCGGTGTTT	TACAGTGGTGTAGTGAGGACCA	6480	7501	CACAGAGAAAGTCTGTGATGTATTT	TAAATGGAATGTCCCTCGGTCCCATCTCCGG	7560
Db	61806	ATTAAACTTTTATTAACAGATTCAGAGGCGGTGTTT	TACAGTGGTGTAGTGAGGACCA	61865	62886	CACAGAGAAAGTCTGTGATGTATTT	TAAATGGAATGTCCCTCGGTCCCATCTCCGG	62945
Qy	6481	TATCTGCCAATGGGAATGGCTGATGAATCCACACTT	TAGTCTCGGATCAGTGTCTGT	6540	7561	GCTCATCGCGGATCTGAATCTCTTT	TAACTGTTCGCTCTACCGGGGTTCCCGGGTTAA	7620
Db	61866	TATCTGCCAATGGGAATGGCTGATGAATCCACACTT	TAGTCTCGGATCAGTGTCTGT	61925	62946	GCTCATCGCGGATCTGAATCTCTTT	TAACTGTTCGCTCTACCGGGGTTCCCGGGTTAA	63005
Qy	6541	GTACTTTACTAGTCCGATGAACCTAAATTTGCGGTGTTT	TAGTCTCGATCAACCAAA	6600	7621	AACCCGGGGGCCGCCACCTGTAACTGTCGCGTGCAGAAATGCGCCAGGGCATCTGTGAG	7680	
Db	61926	GTACTTTACTAGTCCGATGAACCTAAATTTGCGGTGTTT	TAGTCTCGATCAACCAAA	61985	63006	AACCCGGGGGCCGCCACCTGTAACTGTCGCGTGCAGAAATGCGCCAGGGCATCTGTGAG	63065	
Qy	6601	CTGCTCGCCTCGCTTTT	AGAAAGCTCTATGCAAGCTTTAATCGTGTCCAAATAGCTTGT	6660	7691	AATTCTCACCGAGAGGCGCTTTGTTT	TACCGAAGATGTTTATAGCGTCTGCTGCAG	7740
Db	61986	CTGCTCGCCTCGCTTTT	AGAAAGCTCTATGCAAGCTTTAATCGTGTCCAAATAGCTTGT	62045	63066	AATTCTCACCGAGAGGCGCTTTGTTT	TACCGAAGATGTTTATAGCGTCTGCTGCAG	63125
Qy	6661	TTTAAATGGAATCTGGTACATCTTTATATGACAGTGTCTCAAAAACCAAGATTTAA	6720	62046	7741	CGGGTCTGTTATACCGGCCACAGTTGGCCAGAGTTCTGCACGACGTCTACCGGAAATGAA	7800	
Db	62046	TTTAAATGGAATCTGGTACATCTTTATATGACAGTGTCTCAAAAACCAAGATTTAA	62105	62106	63126	CGGGTCTGTTATACCGGCCACAGTTGGCCAGAGTTCTGCACGACGTCTACCGGAAATGAA	63185	
Qy	6721	AAACAGCTTAACTCGGTGCTGCTGTAGTCTCTGATATATACTCTCAACCAAGAA	6780	62165	7801	GGCCAAAGTGTGGGGGCTGGCGTCTCTCATATGTTGAGGGCGCTATCATGGCTAT	7860	
Db	62106	AAACAGCTTAACTCGGTGCTGCTGTAGTCTCTGATATATACTCTCAACCAAGAA	62165	6781	63186	GGCCAAAGTGTGGGGGCTGGCGTCTCTCATATGTTGAGGGCGCTATCATGGCTAT	63245	
Qy	6781	TTTTTTAACCAACTCGCCAGGTACTGAAACGATGTTACGGA	CAGTCTGTGAAAGTGT	6840	7861	TGCGACTCGCTCTGCTCACTTATAATACCTCGATGCGAGGAGAACTTGAACCTTAG	7920	
Db	62166	TTTTTTAACCAACTCGCCAGGTACTGAAACGATGTTACGGA	CAGTCTGTGAAAGTGT	62225	63246	TGCGACTCGCTCTGCTCACTTATAATACCTCGATGCGAGGAGAACTTGAACCTTAG	63305	
Qy	6841	TATCATCTCTTACCTCCCTTTGAAAGTGTGTTT	TACACCGACATGTGATAGGCA	6900	7921	ACTCAAGCATCTGTGCAAACTTTT	TCAACCATCTTTCTTCAAGAACTTTAGCCCC	7980
Db	62226	TATCATCTCTTACCTCCCTTTGAAAGTGTGTTT	TACACCGACATGTGATAGGCA	62285	63306	ACTCAAGCATCTGTGCAAACTTTT	TCAACCATCTTTCTTCAAGAACTTTAGCCCC	63365
Qy	6901	CCAGCTCAGGGGGAGTTCGATCGTGAAGGGGTAA	CAATTCGTGGCGTGGGGTACT	6960	7981	TATGAGCTACTTGCCCAACGAGGAGATGTTT	CTTCAACATTTATTCATCAACCGGTAC	8040
Db	62286	CCAGCTCAGGGGGAGTTCGATCGTGAAGGGGTAA	CAATTCGTGGCGTGGGGTACT	62345	63366	TATGAGCTACTTGCCCAACGAGGAGATGTTT	CTTCAACATTTATTCATCAACCGGTAC	63425
Qy	6961	CAGTTTATGATCTCCGCGAGGACGCTAGAGTTT	CTTCTCAATCCTACGACCGCCG	7020	8041	GGCCGAGAGAGGCGCCCTGCTGTGATCTGGAAGTACTAGCGGGATGACGTGTCCGGGAAG	8100	
Db	62346	CAGTTTATGATCTCCGCGAGGACGCTAGAGTTT	CTTCTCAATCCTACGACCGCCG	62405	63426	GGCCGAGAGAGGCGCCCTGCTGTGATCTGGAAGTACTAGCGGGATGACGTGTCCGGGAAG	63485	
Qy	7021	ATCCACAGGGGAGTACTAGTTTAAAGTCCCGGAATCTAT	CTGCTGTGTAATTTTGTG	7080	8101	CAGCCTTGTCCCGATCTCTTAATCTCCGAGTATCGAACCGGGCTTCTCCCGGACCC	8160	
Db	62406	ATCCACAGGGGAGTACTAGTTTAAAGTCCCGGAATCTAT	CTGCTGTGTAATTTTGTG	62465	63486	CAGCCTTGTCCCGATCTCTTAATCTCCGAGTATCGAACCGGGCTTCTCCCGGACCC	63545	
Qy	7081	GAGCTCTTTAGACAGTAAACATTTGCTCCGACACTGCGGTACGATGACCGCTCTCC		7140	8161	CCTGCTGACCTCAGTAAATGTCTTAGAAAAATCCAGAAATCATCTCTCAGCGCCCCACCCCT	8220	
Db	62466	GAGCTCTTTAGACAGTAAACATTTGCTCCGACACTGCGGTACGATGACCGCTCTCC		62525	63546	CCTGCTGACCTCAGTAAATGTCTTAGAAAAATCCAGAAATCATCTCTCAGCGCCCCACCCCT	63605	
Qy	7141	TTGCTGTGTTGATGACACCTCGAGTATCTTAGACAGGTATAGGGCCCGTTCCACTCT		7200	8221	GAGTCAATTTGTTCATCAAAACACGACCCAGTCTGCTCAGTCACTCAGTCACTATTTAC	8280	
Db	62526	TTGCTGTGTTGATGACACCTCGAGTATCTTAGACAGGTATAGGGCCCGTTCCACTCT		62585	63606	GAGTCAATTTGTTCATCAAAACACGACCCAGTCTGCTCAGTCACTCAGTCACTATTTAC	63665	
Qy	7201	AAGGTTCCCGCTTGGTAGCGGATTCGCGGCAATTCAGCGT	CAGTGTTCATTTCAAC	7260	8281	GCCAAACCCAGGGCGTTGTTCCCGGCAATGTTTATGAGACAGTGGAAAGCGGTGTACA	8340	
Db	62586	AAGGTTCCCGCTTGGTAGCGGATTCGCGGCAATTCAGCGT	CAGTGTTCATTTCAAC	62645	63666	GCCAAACCCAGGGCGTTGTTCCCGGCAATGTTTATGAGACAGTGGAAAGCGGTGTACA	63725	
Qy	7261	GACTATCATGCGAGTGCAGATAGTGAATCTTGATCATATGATTTAAAGATGCTCTT		7320	8341	GAGCATTTACCAACAGGACAGACGCTTATTTT	TGCGCGCCGACATAACCGGTTCCGACATC	8400
Db	62646	GACTATCATGCGAGTGCAGATAGTGAATCTTGATCATATGATTTAAAGATGCTCTT		62705	63726	GAGCATTTACCAACAGGACAGACGCTTATTT	TGCGCGCCGACATAACCGGTTCCGACATC	63785
Qy	7321	CCCGTAATGGGACTCATGCTGTGTTTAAAGTTCAGGGTGT	TAAACCTCAGAAG	7380	8401	TSCGGCGCTGCGCGCATATCGCATGTTCCCGAGTTCGCGGACAGTCTCTCGGAGGTGA	8460	
Db	62706	CCCGTAATGGGACTCATGCTGTGTTTAAAGTTCAGGGTGT	TAAACCTCAGAAG	62765	63786	TGCGGCGCTGCGCGCATATCGCATGTTCCCGAGTTCGCGGACAGTCTCTCGGAGGTGA	63845	
Qy	7381	GTCCCTGCGTGAACATTTTTCGAGTCTCCGACCGCTCAGGAGT	TAGCTGTACACGC	7440	8461	AGGGTCCGCGGCGTCTGATTTGCGGAAACACAGCATGCTTCCGAGCGCGGTACCGCC	8520	
Db	62766	GTCCCTGCGTGAACATTTTTCGAGTCTCCGACCGCTCAGGAGT	TAGCTGTACACGC	62825	63846	AGGGTCCGCGGCGTCTGATTTGCGGAAACACAGCATGCTTCCGAGCGCGGTACCGCC	63905	
					8521	CAATGTCCCGGACCAAGAGGATGAAACGGTAGCACAGTTGGGAAACGCTCCCGTAAAAA	8580	

63906 CAATGTCCTCCGCGCAAGAGAGTGAAGACGGTAGCACAGTTGGGAAACGCTCCGCTAAATAA 63965  
8581 CGTCCACATCGGAGCGCGGTATACGCTCCACCTGTTATATACCAATATAGCTTAAC 8640  
63966 CGTCCACATCGGAGCGCGGTATACGCTCCACCTGTTATATACCAATATAGCTTAAC 64025  
8641 GTCCCTCCCTCAGGCTCGGCGCAGAGTCGGCCGATATCGCCAACTCCAGAGTCCCGCAT 8700  
64026 GTCCCTCCCTCAGGCTCGGCGCAGAGTCGGCCGATATCGCCAACTCCAGAGTCCCGCAT 64085  
8701 GCGCGCGCGCTCTCCGCGCTTCGCGCAAAACCGCGCAACGCTCCCGCTAGAGAAAGCA 8760  
64086 GCGCGCGCGCTCTCCGCGCTTCGCGCAAAACCGCGCAACGCTCCCGCTAGAGAAAGCA 64145  
8761 GCGCGCGAGGAGCTGCGCAGACAAAGAGTGAAGGCGCGAGCTTCGCGCGCGCGCAACAGT 8820  
64146 GCGCGCGAGGAGCTGCGCAGACAAAGAGTGAAGGCGCGAGCTTCGCGCGCGCGCAACAGT 64205  
8821 AAACCCACCTTCCTCCGCGCGCTTCGCGCAAAACCGCGCAACGCTCCCGCTAGAGAAAGT 8880  
64206 AAACCCACCTTCCTCCGCGCGCTTCGCGCAAAACCGCGCAACGCTCCCGCTAGAGAAAGT 64265  
8881 AATCGAAGCTCCACGAGTGTAAACCGCAACGATCTGGACCCCAAAACGACGACGACAT 8940  
64266 AATCGAAGCTCCACGAGTGTAAACCGCAACGATCTGGACCCCAAAACGACGACGACAT 64325  
8941 GCTAGCGGCTATCTACAGAGCTGTATGAGTACAGTCCCGCGCGCATCTTCCCTCC 9000  
64326 GCTAGCGGCTATCTACAGAGCTGTATGAGTACAGTCCCGCGCGCATCTTCCCTCC 64385  
9001 CTCAGCAACTCGGCAATGAGGAGATATTTCCAGAGGTAGTCCGCCATCTAGCGGCA 9060  
64386 CTCAGCAACTCGGCAATGAGGAGATATTTCCAGAGGTAGTCCGCCATCTAGCGGCA 64445  
9061 CGGATCGCCTTGA 9073  
64446 CGGATCGCCTTGA 64458

RESULT 2  
AAC64754  
ID AAC64754 standard; DNA; 133719 BP.  
XX  
AC AAC64754;  
XX  
DT 28-FEB-2001 (first entry)  
XX  
DE Macaca mulatta rhadinovirus 17577 (RRV) genome sequence SEQ ID NO:1.  
XX  
KW Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;  
KW genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;  
KW IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;  
KW cytostatic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;  
KW lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;  
KW splenomegaly; hypergammaglobulinemia; autoimmune haemolytic anaemia;  
KW ds.  
XX  
OS Macaca mulatta rhadinovirus 17577.  
XX  
PN W0200028040-A2.  
XX  
PD 18-MAY-2000.  
XX  
PF 05-NOV-1999; 99WO-US026260.  
XX  
PR 06-NOV-1998; 98US-0107507P.  
PR 20-NOV-1998; 98US-0109409P.  
XX  
PA (UOYOR-) UNIV OREGON HEALTH SCI.  
XX  
PI Wong SW, Axthelm MK, Searles RP;  
XX

WPI; 2000-376552/32.  
XX New rhesus rhadino virus for producing non-human primate model useful for  
PT testing potential treatments and efficacy of the candidate vaccine for  
PT conditions associated with RRV infection.  
XX  
PS Claim 2; Page 83-122; 141pp; English.  
XX  
CC The present invention describes a novel rhesus macaque rhadinovirus  
CC called macaca mulatta rhadinovirus 17577 (RRV). AAC64754 represents the  
CC RRV genome sequence, and AAB53123 to AAB53205 represent the proteins  
CC encoded by the genome sequence. The present invention also specifically  
CC claims the individual open reading frame (ORF) nucleotide sequences from  
CC the genome which encode the individual proteins, but these sequences are  
CC not given. A non-human animal infected with RRV can be used for testing  
CC the efficacy of drug in the treatment of condition associated with  
CC infection with RRV such as Kaposi's sarcoma, lymphoproliferative  
CC disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly,  
CC hypergammaglobulinemia or autoimmune haemolytic anaemia, by  
CC administering the drug to an immuno-compromised non-human primate  
CC preferably Rhesus macaque monkey obtained by as a result of infection by  
CC Simian Immunodeficiency Virus (SIV). RRV is useful for producing non-  
CC human primate model for testing potential treatments for conditions  
CC associated with RRV infection. It is also useful for testing the efficacy  
CC of the candidate vaccine against RRV infection or conditions associated  
CC with its infection by administering the vaccine to the subject capable of  
CC infection with RRV, inoculating the subject with RRV and observing the  
CC effect of vaccine. AAC64755 to AAC64765 and AAB53205 to AAB53213  
CC represent sequence used in the exemplification of the present invention  
XX  
SQ Sequence 133719 BP; 32746 A; 35648 C; 34521 G; 30804 T; 0 U; 0 Other;  
Query Match 100.0%; Score 9073; DB 3; Length 133719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGAGTTCGTCGATGACCAAAACAGTCGAGCAAAATCCGGGTTCATCTTGAAC 60  
Db 60966 ATGAGAGTTCGTCGATGACCAAAACAGTCGAGCAAAATCCGGGTTCATCTTGAAC 61025  
QY 61 ATGAGCTCCGATGCCAAAGTCAGGGGTGTCGTGATCAGTCAGTCGCTGTCAAAATA 120  
Db 61026 ATGAGCTCCGATGCCAAAGTCAGGGGTGTCGTGATCAGTCAGTCGCTGTCAAAATA 61085  
QY 121 ACTACAGCCACCGGAAATGGGTGGTACGACCTGCTTCGATCCGGGTGAAGATCC 180  
Db 61086 ACTACAGCCACCGGAAATGGGTGGTACGACCTGCTTCGATCCGGGTGAAGATCC 61145  
QY 181 GGGCGCTTCTTCCCGTTTACCGTTTATCTAATTACGGAACTGCTGCTGGGAAAGT 240  
Db 61146 GGGCGCTTCTTCCCGTTTACCGTTTATCTAATTACGGAACTGCTGCTGGGAAAGT 61205  
QY 241 ACCAGCATATCGCCCTGTACCAAAATTTAACTGCTGTATCAGCGGCGGACCAACATA 300  
Db 61206 ACCAGCATATCGCCCTGTACCAAAATTTAACTGCTGTATCAGCGGCGGACCAACATA 61265  
QY 301 GCGGCACAGAACCTTATCGCTCGCTTAAAGAGCTTCTGTCCACGATCTTTCAGGCTTTT 360  
Db 61266 GCGGCACAGAACCTTATCGCTCGCTTAAAGAGCTTCTGTCCACGATCTTTCAGGCTTTT 61325  
QY 361 GCGCTTAAAGAGCGGACATCAATATAGCCGTACAGAAAGCTATCAGACCGGAGCCGTA 420  
Db 61326 GCGCTTAAAGAGCGGACATCAATATAGCCGTACAGAAAGCTATCAGACCGGAGCCGTA 61385  
QY 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAGTATTTGGCGGTATAGTGGACATT 480  
Db 61386 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAGTATTTGGCGGTATAGTGGACATT 61445  
QY 481 ATGAAGAGGTTATGGCGAATAAACCCTATGGCATGTACGGGACTATATCAACGCGAAT 540  
Db 61446 ATGAAGAGGTTATGGCGAATAAACCCTATGGCATGTACGGGACTATATCAACGCGAAT 61505  
QY 541 TTTGAAACCCCTCTCGAGAAATGACCGGACCGTGTATTATGGACTTCCCAATATTATTGAATC 600

Db 61506 TTTGAAACCTCTCGAGATGACCGGACCGTGTATGAGCTTCCCAATATATTGTAATC 61565  
Qy 601 GACGAGCGGGAACCTCTGCTCTTACATACCTACACACCGTGTGTTCTTTACTGGTTC 660  
Db 61566 GACGAGCGGGAACCTCTGCTCTTACATACCTACACACCGTGTGTTCTTTACTGGTTC 61625  
Qy 661 CTAAACAGCTGCGTAATACCTCTTACCGGACGCGGCGGTTCGGTGCATAGTATGC 720  
Db 61626 CTAAACAGCTGCGTAATACCTCTTACCGGACGCGGCGGTTCGGTGCATAGTATGC 61685  
Qy 721 GTCGGTTCGCAACCGCAGACAAACCGGTTCCAGTCAACTTACCAACCGGACGCAAAAG 780  
Db 61686 GTCGGTTCGCAACCGCAGACAAACCGGTTCCAGTCAACTTACCAACCGGACGCAAAAG 61745  
Qy 781 ACGGAGATATGTCGTCGAGAACATCTTAACATTCATGATAGGAGAGGTGCTATCT 840  
Db 61746 ACGGAGATATGTCGTCGAGAACATCTTAACATTCATGATAGGAGAGGTGCTATCT 61805  
Qy 841 GAGTACGTACCTTGGAGAGAACTGGCGCTGTTTATAAACAAGCGCTGCACCGAT 900  
Db 61806 GAGTACGTACCTTGGAGAGAACTGGCGCTGTTTATAAACAAGCGCTGCACCGAT 61865  
Qy 901 CTGCGAGTTCGACACCTCTTAATAATTTAGAGTATTAATCTTCCCAATCTCGACGAGTC 960  
Db 61866 CTGCGAGTTCGACACCTCTTAATAATTTAGAGTATTAATCTTCCCAATCTCGACGAGTC 61925  
Qy 961 ATGAGTACGTACGAGATTTGTCGTTCTTAAAGTAAGATTAAGATCTTTAGAAATAC 1020  
Db 61926 ATGAGTACGTACGAGATTTGTCGTTCTTAAAGTAAGATTAAGATCTTTAGAAATAC 61985  
Qy 1021 ATTGCGTGGACCGACCTTTTGTACATAGCGAGTAAAGCGGTATCTAACAACTA 1080  
Db 61986 ATTGCGTGGACCGACCTTTTGTACATAGCGAGTAAAGCGGTATCTAACAACTA 62045  
Qy 1081 CACACATGTCATACGCTAGGCGGATACCGAGGACACGAGCTCTTTACCTGTCCGCTG 1140  
Db 62046 CACACATGTCATACGCTAGGCGGATACCGAGGACACGAGCTCTTTACCTGTCCGCTG 62105  
Qy 1141 GTGTGCGAGGTGTTTGTGAGCGCTTTGAGGAATACAGCGGCGCTCAACCTCAAC 1200  
Db 62106 GTGTGCGAGGTGTTTGTGAGCGCTTTGAGGAATACAGCGGCGCTCAACCTCAAC 62165  
Qy 1201 CTCACGCTGACGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 1260  
Db 62166 CTCACGCTGACGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 62225  
Qy 1261 GTGACACGAGCATGTCCATAGTCGCCACGGAATCCACTGAACGTTTCGACCCAGTTACC 1320  
Db 62226 GTGACACGAGCATGTCCATAGTCGCCACGGAATCCACTGAACGTTTCGACCCAGTTACC 62285  
Qy 1321 TTTATCACCAAGTTTGTAAACACGACCGTCTTTAAATGGAACCAAAAAATGC 1380  
Db 62286 TTTATCACCAAGTTTGTAAACACGACCGTCTTTAAATGGAACCAAAAAATGC 62345  
Qy 1381 ATATGTGGTTTCAGGTTAGTACTTTGAGTTCAAAGAAATCCTAGACAGCAACTCTTC 1440  
Db 62346 ATATGTGGTTTCAGGTTAGTACTTTGAGTTCAAAGAAATCCTAGACAGCAACTCTTC 62405  
Qy 1441 GTGAGACGCAATTCGCAAGCCGTCGGAATACGTATACGGTTTCTTAAACACGTTACTG 1500  
Db 62406 GTGAGACGCAATTCGCAAGCCGTCGGAATACGTATACGGTTTCTTAAACACGTTACTG 62465  
Qy 1501 TACAACGCGCATGTACTGTTTACGCGTACGCGTGTAGGTACACAGAGAAATATCTG 1560  
Db 62466 TACAACGCGCATGTACTGTTTACGCGTACGCGTGTAGGTACACAGAGAAATATCTG 62525  
Qy 1561 CAAGACCTAAATTTGACCCCTCCCGCGCTCTGGCAACCGGCGTGTAGACCTTCAA 1620  
Db 62526 CAAGACCTAAATTTGACCCCTCCCGCGCTCTGGCAACCGGCGTGTAGACCTTCAA 62585  
Qy 1621 ACGGTTGCGTGAAGATTTAAACCTTGAAGACGACATCTTTACACCGTGTGAGTCCCCG 1680

Db 62586 ACGGTTGCGTGAAGATTTAAACCTTGAAGACGACATCTTCTACACGTTGTCAGTCCCCG 62645  
Qy 1681 CCAACCGCGGTATACCTCTCCCTCCAGGTTTGGTGCACACGTAATGCGCCCTAAAGAC 1740  
Db 62646 CCAACCGCGGTATACCTCTCCCTCCAGGTTTGGTGCACACGTAATGCGCCCTAAAGAC 62705  
Qy 1741 GTGTTGCGCTCCAGAAATAAGGTGCGTCTGTTTGGCGGAGTGTGAGAGGAA 1800  
Db 62706 GTGTTGCGCTCCAGAAATAAGGTGCGTCTGTTTGGCGGAGTGTGAGAGGAA 62765  
Qy 1801 ACSTTTTCCGCGTTTACCGTTAACATGTCGTAAGGACGAGATTGACTTTGTCCTCCCT 1860  
Db 62766 ACSTTTTCCGCGTTTACCGTTAACATGTCGTAAGGACGAGATTGACTTTGTCCTCCCT 62825  
Qy 1861 TCAGAACGCTCTCAACCGGCTGTTGGCGTTTGCATCGACCGTGTGAATCGTATATAAATTAA 1920  
Db 62826 TCAGAACGCTCTCAACCGGCTGTTGGCGTTTGCATCGACCGTGTGAATCGTATATAAATTAA 62885  
Qy 1921 GGTTACACGTTTTCACCGTAGCGTCTGTCAGGTTTCCACTCTAGTATGATAC 1980  
Db 62886 GGTTACACGTTTTCACCGTAGCGTCTGTCAGGTTTCCACTCTAGTATGATAC 62945  
Qy 1981 CTGAGAAAGAGATGCGCTCCCTCCCTGTCGAGGACTCTAGCGGTTTATCGCGTGCCTA 2040  
Db 62946 CTGAGAAAGAGATGCGCTCCCTCCCTGTCGTCGAGGACTCTAGCGGTTTATCGCGTGCCTA 63005  
Qy 2041 GAGAAATAACATTAACCAAAATGACCGAAACCATGAGGACGCGGAGCGTTTCCAAAGTGC 2100  
Db 63006 GAGAAATAACATTAACCAAAATGACCGAAACCATGAGGACGCGGAGCGTTTCCAAAGTGC 63065  
Qy 2101 TGTGCGGCGGACTATGCGGTCAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 2160  
Db 63066 TGTGCGGCGGACTATGCGGTCAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 63125  
Qy 2161 ATGTCGTTGAGGAGTACCGTGTAGTATTTGGTCCCAACAGAGCTCGACAGACAC 2220  
Db 63126 ATGTCGTTGAGGAGTACCGTGTAGTATTTGGTCCCAACAGAGCTCGACAGACAC 63185  
Qy 2221 GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTATTTGGTCAATGACAGCAACCC 2280  
Db 63186 GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTATTTGGTCAATGACAGCAACCC 63245  
Qy 2281 CTTTAAACCTCTCAGAGAACGAGTCAATACCTCGCCCAAGCATATAGTCCGCGC 2340  
Db 63246 CTTTAAACCTCTCAGAGAACGAGTCAATACCTCGCCCAAGCATATAGTCCGCGC 63305  
Qy 2341 CTCACAAACCCCAACCAACCTCTACTATAATAAATAAATAAATAAATAAATAAATAAATAA 2400  
Db 63306 CTCACAAACCCCAACCAACCTCTACTATAATAAATAAATAAATAAATAAATAAATAAATAA 63365  
Qy 2401 ATTGTGTTTATTTAGTCCAAACGAGGACCGGATAGTGTCTATTTTCCACACACCGGG 2460  
Db 63366 ATTGTGTTTATTTAGTCCAAACGAGGACCGGATAGTGTCTATTTTCCACACACCGGG 63425  
Qy 2461 GGGTGGCCGCTATGTTTACGACCAAGAAACGCGCGGCTCTGCGGTTAGAGTAGGCC 2520  
Db 63426 GGGTGGCCGCTATGTTTACGACCAAGAAACGCGCGGCTCTGCGGTTAGAGTAGGCC 63485  
Qy 2521 CGCCCCCGTGTGAGGTCTATGCTCTTTAGGCGCTGAGTGTGACCAATATAATCA 2580  
Db 63486 CGCCCCCGTGTGAGGTCTATGCTCTTTAGGCGCTGAGTGTGACCAATATAATCA 63545  
Qy 2581 CGTGTCTATCAGCTGTGAGTCTCTGCGTGAACGCGGCTCTTTGACGCGGACAGAAAC 2640  
Db 63546 CGTGTCTATCAGCTGTGAGTCTCTGCGTGAACGCGGCTCTTTGACGCGGACAGAAAC 63605  
Qy 2641 CGCGGCTGTGAGGTATCTGATCTGTCGTCGTCTGAGAGTGAACACAGGTCAATCA 2700  
Db 63606 CGCGGCTGTGAGGTATCTGATCTGTCGTCGTCTGAGAGTGAACACAGGTCAATCA 63665  
Qy 2701 CATCAGTGGCGCGCTGCTCATCGGCTCTGATTCTGTCCTCCATGCTCTCAACAGAGCTC 2760  
Db 63666 CATCAGTGGCGCGCTGCTCATCGGCTCTGATTCTGTCCTCCATGCTCTCAACAGAGCTC 63725



QY 2761 GGTGCTCCACTTCGTGCTCATCTGAAATAAATAGTCTCTGTGCTCTGAGTGTATCCAG 2820  
 Db 63726 GGTGCTCCACTTCGTGCTCATCTGAAATAAATAGTCTCTGTGCTCTGAGTGTATCCAG 63785  
 QY 2821 CAAACAAGGTTCTTCTTCATCGGAAGACCATCCGCTCTGCGGTGATCTAGATATAT 2880  
 Db 63786 CAAACAAGGTTCTTCTTCATCGGAAGACCATCCGCTCTGCGGTGATCTAGATATAT 63845  
 QY 2881 CACTGACCAAGCGGCGAGGCGCGTGGGGAAGCGGTGAGTCTCTCGGCACCTTTCAT 2940  
 Db 63846 CACTGACCAAGCGGCGAGGCGCGTGGGGAAGCGGTGAGTCTCTCGGCACCTTTCAT 63905  
 QY 2941 CAGTGTCTCTGAGTATCTATAAAAAATCTCTTCGCTGCTGATTCAGATGACTGGG 3000  
 Db 63906 CAGTGTCTCTGAGTATCTATAAAAAATCTCTTCGCTGCTGATTCAGATGACTGGG 63965  
 QY 3001 GTGCGCGGTTGGCGAGTAACACCGCAACGGCGCGGACCAACCAACGGCGTCACTG 3060  
 Db 63966 GTGCGCGGTTGGCGAGTAACACCGCAACGGCGCGGACCAACCAACGGCGTCACTG 64025  
 QY 3061 GCGCGGCGCTCGGTGCTGTTGGTGGCACCTGTGCAAGTAATGACGTAAGTACCGGC 3120  
 Db 64026 GCGCGGCGCTCGGTGCTGTTGGTGGCACCTGTGCAAGTAATGACGTAAGTACCGGC 64085  
 QY 3121 CGCTCTGTTAAGGACTCACCGTCCGCTGCGTGGATCTATTCTAGTGGGAGGAGA 3180  
 Db 64086 CGCTCTGTTAAGGACTCACCGTCCGCTGCGTGGATCTATTCTAGTGGGAGGAGA 64145  
 QY 3181 CCGCGGTTCTCTTCTTCCATCCATCACTAGTCTGCTGCTCCGATGAGAGTGG 3240  
 Db 64146 CCGCGGTTCTCTTCTTCCATCCATCACTAGTCTGCTGCTCCGATGAGAGTGG 64205  
 QY 3241 TTTCTGCTCGGTTTGTGAGTAGTGGAGGGGATCTTGGTCAAAACATCTCAGTGG 3300  
 Db 64206 TTTCTGCTCGGTTTGTGAGTAGTGGAGGGGATCTTGGTCAAAACATCTCAGTGG 64265  
 QY 3301 CAGTCACTGCAATCACTGCGGCGCTCCGTACCGGCGGTACCGCAACCGGTACCGG 3360  
 Db 64266 CAGTCACTGCAATCACTGCGGCGCTCCGTACCGGCGGTACCGCAACCGGTACCGG 64325  
 QY 3361 ACCCGCGAACAAAAACGCTCTGTTCTCTTTTCTAGTGCCTGCGGGAATCGCAGCA 3420  
 Db 64326 ACCCGCGAACAAAAACGCTCTGTTCTCTTTTCTAGTGCCTGCGGGAATCGCAGCA 64385  
 QY 3421 TCTCGGAGTGGCGCGGCGCTTTCAGCGGTGCGATAGCAATAGCCATGCGCGGAA 3480  
 Db 64386 TCTCGGAGTGGCGCGGCGCTTTCAGCGGTGCGATAGCAATAGCCATGCGCGGAA 64445  
 QY 3481 CGTTCACTGTAAACGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGCGCAC 3540  
 Db 64446 CGTTCACTGTAAACGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGCGCAC 64505  
 QY 3541 CTTCCGATTCGAACTAGATAGTCAATCGCAAGTTTAAATGGTTGCGAGCTAGAACTG 3600  
 Db 64506 CTTCCGATTCGAACTAGATAGTCAATCGCAAGTTTAAATGGTTGCGAGCTAGAACTG 64565  
 QY 3601 CGGCCAGGAGAACCGGTTGCGCGTGGCGGTAGCCAGGAGACGGATGCGCGCTT 3660  
 Db 64566 CGGCCAGGAGAACCGGTTGCGCGTGGCGGTAGCCAGGAGACGGATGCGCGCTT 64625  
 QY 3661 GAGGACGAGATGCTCTGCTCGTTTATGACGCTGCGCTTGTCTATAGCTTTCTCCCA 3720  
 Db 64626 GAGGACGAGATGCTCTGCTCGTTTATGACGCTGCGCTTGTCTATAGCTTTCTCCCA 64685  
 QY 3721 CAGCATAAAGCGACTCTGAGCTTGGCAGAAAGCGACTTATATGCTAGCTGTGAA 3780  
 Db 64686 CAGCATAAAGCGACTCTGAGCTTGGCAGAAAGCGACTTATATGCTAGCTGTGAA 64745  
 QY 3781 CCAATCCAGCGAGGTTGCGGTGTGACCCCGCTTCCCTCTTTCCACCGTCAAGATGGT 3840  
 Db 64746 CCAATCCAGCGAGGTTGCGGTGTGACCCCGCTTCCCTCTTTCCACCGTCAAGATGGT 64805

QY 3841 GTTTAAAGCAGAACTCCCGTTTTGCCCACAAAGTCTTAAGCACCCGTGAGAGGAGCGGT 3900  
 Db 64806 GTTTAAAGCAGAACTCCCGTTTTGCCCACAAAGTCTTAAGCACCCGTGAGAGGAGCGGT 64865  
 QY 3901 GAACCCAGGTACAGTATGCTCTCTTTTAAATAATTTTGGAGGCTTGGAGGTATCT 3960  
 Db 64866 GAACCCAGGTACAGTATGCTCTCTTTTAAATAATTTTGGAGGCTTGGAGGTATCT 64925  
 QY 3961 GTAGTCCGGAGCAACGCTAAAGCTAGTCCGGTGTGACCCGCGATGGTAAGATCCTG 4020  
 Db 64926 GTAGTCCGGAGCAACGCTAAAGCTAGTCCGGTGTGACCCGCGATGGTAAGATCCTG 64985  
 QY 4021 GCCAAGATCAGACCTTTAATATCCTCCGTTCCGCAACATACGACCCACACATAATCT 4080  
 Db 64986 GCCAAGATCAGACCTTTAATATCCTCCGTTCCGCAACATACGACCCACACATAATCT 65045  
 QY 4081 ATCTATCCGTGGGTAATATCAGGTTAGTCTCATATCCATTAACCGCTTCCAGAGC 4140  
 Db 65046 ATCTATCCGTGGGTAATATCAGGTTAGTCTCATATCCATTAACCGCTTCCAGAGC 65105  
 QY 4141 CGGAGCTTCTGTTTAAATAATCGAAGATTTAAATAATCAAGCCATGAGTCCGTAAG 4200  
 Db 65106 CGGAGCTTCTGTTTAAATAATCGAAGATTTAAATAATCAAGCCATGAGTCCGTAAG 65165  
 QY 4201 TAGCAAGCTTGTGAGGAGTCTTCTTAACACTTCTGTTGACATTTTACTCCAAACTAT 4260  
 Db 65166 TAGCAAGCTTGTGAGGAGTCTTCTTAACACTTCTGTTGACATTTTACTCCAAACTAT 65225  
 QY 4261 TGTTTAAGCAACCTTCCATAATCCCGCGTCTGTTAATTTACTTACGTTACGTTAA 4320  
 Db 65226 TGTTTAAGCAACCTTCCATAATCCCGCGTCTGTTAATTTACTTACGTTACGTTAA 65285  
 QY 4321 ACTTTGCTGCTGCTGAGAAATCCACAGAGGTGCGCGGTATTAAGCAGTTCGCTT 4380  
 Db 65286 ACTTTGCTGCTGCTGAGAAATCCACAGAGGTGCGCGGTATTAAGCAGTTCGCTT 65345  
 QY 4381 TCACTACCGTTTATGTAAGTCTTCAACTCGGCTTAAACATCATCAAGACTATATTCG 4440  
 Db 65346 TCACTACCGTTTATGTAAGTCTTCAACTCGGCTTAAACATCATCAAGACTATATTCG 65405  
 QY 4441 ATATAATTCACAAATTCAGCTCTTCTCGGAGCACTGTAATTTAACTTGTACAAATC 4500  
 Db 65406 ATATAATTCACAAATTCAGCTCTTCTCGGAGCACTGTAATTTAACTTGTACAAATC 65465  
 QY 4501 CTTTAAATAAACGCTTAAAGTGAAATCCATTTACACATATTTCTGTTGAGAGCACTA 4560  
 Db 65466 CTTTAAATAAACGCTTAAAGTGAAATCCATTTACACATATTTCTGTTGAGAGCACTA 65525  
 QY 4561 GCTTTATATCGCAATGTTGCTGTACACAGCCCTGGACATAAAGCTAAATTTGTTAAATGA 4620  
 Db 65526 GCTTTATATCGCAATGTTGCTGTACACAGCCCTGGACATAAAGCTAAATTTGTTAAATGA 65585  
 QY 4621 AAAGAAGCAATCTCAACCGCTCCGGAATATAAAGTCACTAAGTGTATTTTACAA 4680  
 Db 65586 AAAGAAGCAATCTCAACCGCTCCGGAATATAAAGTCACTAAGTGTATTTTACAA 65645  
 QY 4681 CATGGATTTACTAAGCCCGGTACCAATGTAATTTAAATAAATCCACATGTAATAAT 4740  
 Db 65646 CATGGATTTACTAAGCCCGGTACCAATGTAATTTAAATAAATCCACATGTAATAAT 65705  
 QY 4741 AAACCTTAAGGTATACATCTTCTCATGTTTATACATAATATGAACGCTAATGCTGACT 4800  
 Db 65706 AAACCTTAAGGTATACATCTTCTCATGTTTATACATAATATGAACGCTAATGCTGACT 65765  
 QY 4801 TATATAATGTATAGAAATTAATGCCAATACAGTCAATGTAAGTATATTTTAAGGCAATAAT 4860  
 Db 65766 TATATAATGTATAGAAATTAATGCCAATACAGTCAATGTAAGTATATTTTAAGGCAATAAT 65825  
 QY 4861 GCAACATCAAAATATATACACAAATGACGCTCCGGATATCCGCCAGCTGTTAAC 4920  
 Db 65826 GCAACATCAAAATATATACACAAATGACGCTCCGGATATCCGCCAGCTGTTAAC 65885  
 QY 4921 CCGAAAATACAGAAATGACTACACAAAACACACCTGAAACCAATTTTATCTCTCAACATATG 4980



65886	CGAAATACAGAAATGACTACACAAACACACTCGAAACCAATTTATTTCTCAACATATG	65945
4981	CAACAATATTTCCAGGTACACCATGTTTAATAAATACGCGACATGCACATTTTAGCTA	5040
65946	CAACAATATTTCCAGGTAAACACCATGTTAATAAATACGCGACATGCACATTTTAGCTA	66005
5041	AGCTCCTCAAGACAAATTTCCCTTTCAATTTGATGACTCATCAACCTCGGTTTCGTTGTTA	5100
66006	AGCTCCTCAAGACAAATTTCCCTTTCAATTTGATGACTCATCAACCTCGGTTTCGTTGTTA	66065
5101	ATGTCGGAAATATATGAGTCCAGGATACCTCGTGCATCGCCACAATTAATTCAGGGCT	5160
66066	ATGTCGGAAATATATGAGTCCAGGATACCTCGTGCATCGCCACAATTAATTCAGGGCT	66125
5161	TGTAAATATCATCCAGTCAATCAGCTGCTAATGTCAAGATGTTGTTCTTAATTTTACG	5220
66126	TGTAAATATCATCCAGTCAATCAGCTGCTAATGTCAAGATGTTGTTCTTAATTTTACG	66185
5221	CATCCAAATGTGAGTGGGTGACACTTTGGACACAGACCCCGTAAAAATTTGGCAGCTCC	5280
66186	CCATCCAAATGTGAGTGGGTGACACTTTGGACACAGACCCCGTAAAAATTTGGCAGCTCC	66245
5281	ACACAAACGGAAGGATGCAATTTGATCGTTTCAGAGGGGAGCATATTTATTTTATAACA	5340
66246	ACACAAACGGAAGGATGCAATTTGATCGTTTCAGAGGGGAGCATATTTATTTTATAACA	66305
5341	TACGAAACCGGACCCGTCAATGACGACATATGTCATTTAAGATTTGGAGTCGTGACACTG	5400
66306	TACGAAACCGGACCCGTCAATGACGACATATGTCATTTAAGATTTGGAGTCGTGACACTG	66365
5401	GTTGAATTCAGCGTGAAGAAATATGTTTCGTGTCACATCGCTGAATACTTTTAAACAC	5460
66366	GTTGAATTCAGCGTGAAGAAATATGTTTCGTGTCACATCGCTGAATACTTTTAAACAC	66425
5461	ACGAGGGCGGAATAAAACGGGTCCCGCGCTTAAGCCGTGAATCGCGCAGCACGGCTC	5520
66426	ACGAGGGCGGAATAAAACGGGTCCCGCGCGCTTAAGCCGTGAATCGCGCAGCACGGCTC	66485
5521	ATTGCGCATTTCCAGATTTGTTAAATCTGTTTCGAAATACGAGTCTAAATCAAGTATAAC	5580
66486	ATTGCGCATTTCCAGATTTGTTAAATCTGTTTCGAAATACGAGTCTAAATCAAGTATAAC	66545
5581	CACGACAAGTTCAACAGGAGGCAACCGCGCTGGTTTCCACCGGACATCCCGCTCAACA	5640
66546	CACGACAAGTTCAACAGGAGGCAACCGCGCTGGTTTCCACCGGACATCCCGCTCAACA	66605
5641	AACCCGAACTCTCTAGGAAACGACATCAAGGGTTTCAGACAAAATTCGGAACCGAACC	5700
66606	AACCCGAACTCTCTAGGAAACGACATCAAGGGTTTCAGACAAAATTCGGAACCGAACC	66665
5701	CTAGTTCTGAGCAACCATCTGGTTATATAAATCGCGGTTAAGAGACGAGCGGCCACACTGC	5760
66666	CTAGTTCTGAGCAACCATCTGGTTATATAAATCGCGGTTAAGAGACGAGCGGCCACACTGC	66725
5761	TGCTTGTAGGAATTCAGCTTAAGTGGCTGCAGGCTTGCTGCACAGGCCGCTAGCTGCCTG	5820
66726	TGCTTGTAGGAATTCAGCTTAAGTGGCTGCAGGCTTGCTGCACAGGCCGCTAGCTGCCTG	66785
5821	GCGGCGCTCCGACACTTTTGACAGCGCACGCCAGTTCCAGAGCGAGGACAGAGCGCTCT	5880
66786	GCGGCGCTCCGACACTTTTGACAGCGCACGCCAGTTCCAGAGCGAGGACAGAGCGCTCT	66845
5881	TTGAGTTCGGCCATATCTCGGCCATAGAGATTCGCTCGCGTGAATCTTTTAGTGTGATA	5940
66846	TTGAGTTCGGCCATATCTCGGCCATAGAGATTCGCTCGCGTGAATCTTTTAGTGTGATA	66905
5941	TATATCTGTTGGGTAAACAACTTCCTCCCTGTAAGGGGAATAAGGTCGTTCTTTCT	6000
66906	TATATCTGTTGGGTAAACAACTTCCTCCCTGTAAGGGGAATAAGGTCGTTCTTTCT	66965
6001	ATCGTTTCTCTGAAATCTATAGTTACATCTTTAAGTATAACCAAGACGCTAAAAACCA	6060

Db	6966	ATCGTTTCTCTGAATCTATAGTTTACATCTTTTAAGTATPAACAAAGACGCTTAAACCCCA	67025
Qy	6061	GGATGGTCGACACACCTTTTAAAAATCGCAAAAGCGACCGATGGCTCTATCGCGTTTCCA	6120
Db	67026	GGATGGTCGACACCTTTTAAAAATCGCAAAAGCGACCGATGGCTCTATCGCGTTTCCA	67085
Qy	6121	TGTCGTTTCAAAGGTAGTACGATGATCTCAGATTAGACTCTGTCTCCCTGTTTAAACACC	6180
Db	67086	TGTCGTTTCAAAGGTAGTACGATGATCTCAGATTAGACTCTGTCTCCCTGTTTAAACACC	67145
Qy	6181	TTAACTGGGAATGAAACAGCCATGACACCGTGAACTTCCTGATGTCCTCTAAAAACTAAAC	6240
Db	67146	TTAACTGGGAATGAAACAGCCATGACACCGTGAACTTCCTGATGTCCTCTAAAAACTAAAC	67205
Qy	6241	CCGAAACAGAGCTTAATACCAANGACTGTCAACCCCTACCCCAAGCAGCGCCCCGTTACTA	6300
Db	67206	CCGAAACAGAGCTTAATACCAANGACTGTCAACCCCTACCCCAAGCAGCGCCCCGTTACTA	67265
Qy	6301	TTAGACCGGGTGAGTAACCAACGCTATCCTTTAAAAACCCATACGTCGAGTTTCTTAAGGT	6360
Db	67266	TTAGACCGGGTGAGTAACCAACGCTATCCTTTAAAAACCCATACGTCGAGTTTCTTAAGGT	67325
Qy	6361	AAACAGCTCGTGATATTTTACAGACCCCTGT CATAAATGGATACCGACGACAATCAGGTA	6420
Db	67326	AAACAGCTCGTGATATTTTACAGACCCCTGT CATAAATGGATACCGACGACAATCAGGTA	67385
Qy	6421	ATTAAACTTTTTTATCTCAAGATT CAGAGGGCGTGTTTACAGTGTGTAGTGGGAGCA	6480
Db	67386	ATTAAACTTTTTTATCTCAAGATT CAGAGGGCGTGTTTACAGTGTGTAGTGGGAGCA	67445
Qy	6481	TATCTCGCCAAATGGGAATGGCTGATGAATCCACACTTAGTCTCGGATCAGGTGCTCTGT	6540
Db	67446	TATCTCGCCAAATGGGAATGGCTGATGAATCCACACTTAGTCTCGGATCAGGTGCTCTGT	67505
Qy	6541	GTACTTTATPAGTGCAGTGAACCTTAATTTGGCGTGTTTAGTGTCTACCAACCAAAAA	6600
Db	67506	GTACTTTATPAGTGCAGTGAACCTTAATTTGGCGTGTTTAGTGTCTACCAACCAAAAA	67565
Qy	6601	CTGCTGGCCCTGCCCTTTTAGAAGCTCATGCAAGCTTTAATCGTGTCCATAGCTGTCTC	6660
Db	67566	CTGCTGGCCCTGCCCTTTTAGAAGCTCATGCAAGCTTTAATCGTGTCCATAGCTGTCTC	67625
Qy	6661	TTTAAAAATGCAACTCTGGTATCATCTTTATGACAGTGTGTCAAAAAACAAGATTAA	6720
Db	67626	TTTAAAAATGCAACTCTGGTATCATCTTTATGACAGTGTGTCAAAAAACAAGATTAA	67685
Qy	6721	AAACAGCTTAAACTGGTGTCCTGGTGTCTCTGATATAACCTCCTCAACAGAAAAAA	6780
Db	67686	AAACAGCTTAAACTGGTGTCCTGGTGTCTCTGATATAACCTCCTCAACAGAAAAAA	67745
Qy	6781	TTTTTTTAAACCAACTCGCCAGGTACTGAAAACGATGGTACGACAGGTCTGTGAAAAGTGT	6840
Db	67746	TTTTTTTAAACCAACTCGCCAGGTACTGAAAACGATGGTACGACAGGTCTGTGAAAAGTGT	67805
Qy	6841	TATCATCTCTTCAACTCCCTTGAAGTTTTGGTTACCGACCATGTGAGATAGGCA	6900
Db	67806	TATCATCTCTTCAACTCCCTTGAAGTTTTGGTTACCGACCATGTGAGATAGGCA	67865
Qy	6901	CCAGCTCAGGGGGAGGTCCGATCTGTGAGGGGTAACTTCTGTGGCGTGGGGTACTC	6960
Db	67866	CCAGCTCAGGGGGAGGTCCGATCTGTGAGGGGTAACTTCTGTGGCGTGGGGTACTC	67925
Qy	6961	CAGTTTTAGATCTCCGGCAGGACCGGTCTCTTCTGTCMAATCTTACGACGCCGC	7020
Db	67926	CAGTTTTAGATCTCCGGCAGGACCGGTCTCTTCTGTCMAATCTTACGACGCCGC	67985
Qy	7021	ATCCACACAGGGCAGTACTGAGTTAATAGTCCCCGAAATTTCTATCGTTTCGTAATTTTGTG	7080
Db	67986	ATCCACACAGGGCAGTACTGAGTTAATAGTCCCCGAAATTTCTATCGTTTCGTAATTTTGTG	68045
Qy	7081	GAGCTCTCTTAGACAGTAAACCAATGTTGCTCGACACTGCGGTACGATGAACCGCTCTCC	7140
Db	68046	GAGCTCTCTTAGACAGTAAACCAATGTTGCTCGACACTGCGGTACGATGAACCGCTCTCC	68105

Qy	7141	TTGTCGTGTTGCATGAGCACCTGAGGATCTTTAGACAGGTATAGGGCCCGTTCACCTCT	7200
Ds	68106	TTGTCGTGTTGCATGAGCACCTGAGGATCTTTAGACAGGTATAGGGCCCGTTCACCTCT	68165
Qy	7201	AGGGTTCCCGTTGGTAGAGCGGATTCGGGCGATTCAGGTGATTCATTTCACC	7260
Ds	68166	AGGGTTCCCGTTGGTAGAGCGGATTCGGGCGATTCAGGTGATTCATTTCACC	68225
Qy	7261	GACTATCATGCCAGTGGCGGATAGTGAACCTTGTACATATGATTTAAAGATGGTCTTT	7320
Ds	68226	GACTATCATGCCAGTGGCGGATAGTGAACCTTGTACATATGATTTAAAGATGGTCTTT	68285
Qy	7321	CCCGTAATGCGGACATCATGTCGTGTTTAAAGTTTCAAGGTTGTTAAAGTCAAGAG	7380
Ds	68286	CCCGTAATGCGGACATCATGTCGTGTTTAAAGTTTCAAGGTTGTTAAAGTCAAGAG	68345
Qy	7381	GTCTCGCGTGACAACTTTTGGGACTCTCCGACCGCTCAGGGATAGCTGTACACGC	7440
Ds	68346	GTCTCGCGTGACAACTTTTGGGACTCTCCGACCGCTCAGGGATAGCTGTACACGC	68405
Qy	7441	GGTGTCTTTTAAACGAAGCATAGTCAAAAAATGACCCCATGTAATTAAGCT	7500
Ds	68406	GGTGTCTTTTAAACGAAGCATAGTCAAAAAATGACCCCATGTAATTAAGCT	68465
Qy	7501	CACAGAAGTCTCTCGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCATCTCCGG	7560
Ds	68466	CACAGAAGTCTCTCGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCATCTCCGG	68525
Qy	7561	GCTCATCGCGATCTGAATCTCTTTAACCTGTTCTGCTCTACCGGGTTCCTCCGGTTAA	7620
Ds	68526	GCTCATCGCGATCTGAATCTCTTTAACCTGTTCTGCTCTACCGGGTTCCTCCGGTTAA	68585
Qy	7621	AACCCGGGGCGGCACCTGTAAGTTCGCGTGCAGAAATGCGCCAGGGCATCTGTAG	7680
Ds	68586	AACCCGGGGCGGCACCTGTAAGTTCGCGTGCAGAAATGCGCCAGGGCATCTGTAG	68645
Qy	7681	AATTTCTCACGAGAGGGCCCTTTGTTGACCGAAAGATGTTTATAGCGTCTGCTGCGAG	7740
Ds	68646	AATTTCTCACGAGAGGGCCCTTTGTTGACCGAAAGATGTTTATAGCGTCTGCTGCGAG	68705
Qy	7741	CGGGTCTGTTATACCGCCCGAGTTCGCGAGAGTCTGACAGAGTCTACGGGAAATGAA	7800
Ds	68706	CGGGTCTGTTATACCGCCCGAGTTCGCGAGAGTCTGACAGAGTCTACGGGAAATGAA	68765
Qy	7801	GGCCAAAGTGTGGGGGTGGCGTCTCATATGTTGACGGGGCTTATCATGGCTAT	7860
Ds	68766	GGCCAAAGTGTGGGGGTGGCGTCTCATATGTTGACGGGGCTTATCATGGCTAT	68825
Qy	7861	TGCGGACTCGTCTCGTCACTTATATAACCTTGGATCCGAGGGAACCTGGAACCTAG	7920
Ds	68826	TGCGGACTCGTCTCGTCACTTATATAACCTTGGATCCGAGGGAACCTGGAACCTAG	68885
Qy	7921	ACTCAAGCACTGTCGAACCTGTTTCAACCCATCTTCTTCAAGAAATCTTAGCCCC	7980
Ds	68886	ACTCAAGCACTGTCGAACCTGTTTCAACCCATCTTCTTCAAGAAATCTTAGCCCC	68945
Qy	7981	TATCAGCTACTGGCCCAACGGGAAGATGTTTCTGACAAATTTATTTTACCATCAGCGTAC	8040
Ds	68946	TATCAGCTACTGGCCCAACGGGAAGATGTTTCTGACAAATTTATTTTACCATCAGCGTAC	69005
Qy	8041	GGCCGAGAGAGCGCCCTGTCGTGACTGGAAGTACTAGCGGATGACGTGTCGGGAAG	8100
Ds	69006	GGCCGAGAGAGCGCCCTGTCGTGACTGGAAGTACTAGCGGATGACGTGTCGGGAAG	69065
Qy	8101	CAGCCTTGTCCCGATTCTTAACTCTGACGATGCGAACCGGGGCTTCTCCCGGCACC	8160
Ds	69066	CAGCCTTGTCCCGATTCTTAACTCTGACGATGCGAACCGGGGCTTCTCCCGGCACC	69125
Qy	8161	CTTGTTGACCTCAGTAAATGTTTAAAGAAATCAGAAATCATCTCAGCGGCCCAACCT	8220
Ds	69126	CTTGTTGACCTCAGTAAATGTTTAAAGAAATCAGAAATCATCTCAGCGGCCCAACCT	69185
Qy	8221	GAGTCAATTTGTCTATCAGAAACAGCACCCAGTCTGCCTCAGTCACTCAGCATTTATAC	8280
Ds	69186	GAGTCAATTTGTCTATCAGAAACAGCACCCAGTCTGCCTCAGTCACTCAGCATTTATAC	69245
Qy	8281	GCCAAACCCAGGGGTTGTTCCCGGCCAATGTTTATGGAACAGTGGGAAAGCGGTGTACA	8340
Ds	69246	GCCAAACCCAGGGGTTGTTCCCGGCCAATGTTTATGGAACAGTGGGAAAGCGGTGTACA	69305
Qy	8341	GAGCATTCACCAACAGGCAAGACGCTTATTTTGGCGCGGCACTAACCGGTTTCGACATC	8400
Ds	69306	GAGCATTCACCAACAGGCAAGACGCTTATTTTGGCGCGGCACTAACCGGTTTCGACATC	69365
Qy	8401	TGCGGCCCTTGGGCCCATATCGATGTTTCCCGAGTTCGCGGCACTCTCAGCAGGTGGA	8460
Ds	69366	TGCGGCCCTTGGGCCCATATCGATGTTTCCCGAGTTCGCGGCACTCTCAGCAGGTGGA	69425
Qy	8461	AGGGTCCGGGGCGTTCGATTCGGGAAACAGCATGCTTTCGCGACGCGCGGTTCGCGCC	8520
Ds	69426	AGGGTCCGGGGCGTTCGATTCGGGAAACAGCATGCTTTCGCGACGCGCGGTTCGCGCC	69485
Qy	8521	CAATGTCCTCCGCAAGAGGATGGAAACGGTATGACATGTTTGGGAAACGCTCCGTTAAATA	8580
Ds	69486	CAATGTCCTCCGCAAGAGGATGGAAACGGTATGACATGTTTGGGAAACGCTCCGTTAAATA	69545
Qy	8581	CGTCCCATCGAGGCGCGGTATACGCTCCACTGTTTAAATATATACCAATATAGACTTAAC	8640
Ds	69546	CGTCCCATCGAGGCGCGGTATACGCTCCACTGTTTAAATATATACCAATATAGACTTAAC	69605
Qy	8641	GTCCCGTCAAGGTTCGGCGAGAGTTCGGCGGATATGCGCAACACTCCAGAGTCCCGCAT	8700
Ds	69606	GTCCCGTCAAGGTTCGGCGAGAGTTCGGCGGATATGCGCAACACTCCAGAGTCCCGCAT	69665
Qy	8701	GGCGGCGGCTCTCCGCCCTTCCGCAAAACCGCCGCAACGGTCCCGCTTAAGAGAAAGCA	8760
Ds	69666	GGCGGCGGCTCTCCGCCCTTCCGCAAAACCGCCGCAACGGTCCCGCTTAAGAGAAAGCA	69725
Qy	8761	GCCACGCGAGGACGTGGGACAAAGACTGAAGGGCGAGCTTCGGGGCGCCGACACAT	8820
Ds	69726	GCCACGCGAGGACGTGGGACAAAGACTGAAGGGCGAGCTTCGGGGCGCCGACACAT	69785
Qy	8821	AAACCACTTCCCGGACCGTCCGGATGCGGCTTCGCGAGCAGGGCTTATTCGATTT	8880
Ds	69786	AAACCACTTCCCGGACCGTCCGGATGCGGCTTCGCGAGCAGGGCTTATTCGATTT	69845
Qy	8881	AATCGAAAGTCCACCGGATGTAACCGCAACCGCATCTGGAACCAAAACGACACACAT	8940
Ds	69846	AATCGAAAGTCCACCGGATGTAACCGCAACCGCATCTGGAACCAAAACGACACACAT	69905
Qy	8941	GCTAGCGGCTATCTTACGACCGTCTGATGACTACAGTCCCGCCCGCCCATCGATTCCCC	9000
Ds	69906	GCTAGCGGCTATCTTACGACCGTCTGATGACTACAGTCCCGCCCGCCCATCGATTCCCC	69965
Qy	9001	CTCAGCAACTCGGCAACATGAGGAGATATTTCCAGAGGTAGTCCGCGCATCTAGCGGCCA	9060
Ds	69966	CTCAGCAACTCGGCAACATGAGGAGATATTTCCAGAGGTAGTCCGCGCATCTAGCGGCCA	70025
Qy	9061	CGGATCGCCTTGA 9073	
Ds	70026	CGGATCGCCTTGA 70038	
RESULT 3			
AAV19941			
ID	AAV19941	standard; DNA; 137507 BP.	
XX	AAV19941;		
AC	AC		
XX	27-AUG-2003 (revised)		
DT	03-AUG-1998 (first entry)		
XX	KSHV long unique coding region and terminal repeat.		
DE	KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;		
KW			

KW interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;  
 KW complement-binding protein; glycoprotein; capsid protein IV; infection;  
 KW immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;  
 KW lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;  
 KW HIV immune status; anti-inflammatory agent; therapy; ds.  
 XX  
 OS Human herpesvirus 8.

XX Key Location/Qualifiers  
 FH CDS 1142..2794  
 FT /tag= a  
 FT /product= "complement-binding protein"  
 FT 8699..11236  
 FT /tag= b  
 FT /product= "glycoprotein B"  
 FT complement(17261..17875)  
 FT /tag= c  
 FT /product= "interleukin 6"  
 FT complement(21548..21832)  
 FT /tag= d  
 FT /product= "macrophage inflammatory protein II"  
 FT complement(27137..27424)  
 FT /tag= e  
 FT /product= "interferon regulatory factor 1"  
 FT 28661..29741  
 FT /tag= f  
 FT /product= "protein T1.1"  
 FT complement(58976..60175)  
 FT /tag= g  
 FT /product= "glycoprotein M"  
 FT complement(69412..69915)  
 FT /tag= h  
 FT /product= "glycoprotein L"  
 FT complement(88410..88910)  
 FT /tag= i  
 FT /product= "interferon regulatory factor 2"  
 FT 89600..90541  
 FT /tag= j  
 FT /product= "interferon regulatory factor 3"  
 FT 90173..90643  
 FT /tag= k  
 FT /product= "glycoprotein X"  
 FT complement(93636..94127)  
 FT /tag= l  
 FT /product= "interferon regulatory factor 4"  
 FT complement(111931..112443)  
 FT /tag= m  
 FT /product= "capsid protein IV"  
 FT complement(123808..127296)  
 FT /tag= n  
 FT /product= "immediate early protein"

XX W09804576-A1.

XX 05-FEB-1998.

XX 22-JUL-1997; 97MO-US013346.

XX 25-JUL-1996; 96US-00686243.

XX 25-JUL-1996; 96US-00686349.

XX 25-JUL-1996; 96US-00686350.

XX 25-JUL-1996; 96US-00687253.

XX 05-SEP-1996; 96US-00708678.

XX 10-OCT-1996; 96US-00728323.

XX 13-NOV-1996; 96US-00747887.

XX 29-NOV-1996; 96US-00757659.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Chang Y, Bohenzky RA, Russo JJ, Edelman IS, Moore PS;

XX WPI; 1998-130615/12.

XX

PT New nucleic acid encoding Kaposi's sarcoma associated herpes virus  
 PT proteins - useful for, e.g. detecting levels of HHV8 in, and preparation  
 PT of vaccines for treatment of, HIV patients.  
 XX

PS Example 2; Page 135-203; 230pp; English.

XX This sequence represents the long unique region and terminal repeat of  
 CC the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known  
 CC as human herpes virus 8 (HHV8). This sequence contains the DNAs of the  
 CC invention which encode KSHV polypeptides selected from: (a) viral  
 CC macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);  
 CC (c) viral Irf 1; (d) complement-binding protein; glycoproteins B, M or L;  
 CC (d) capsid protein IV encoded by ORF65; and (e) immediate early protein  
 CC encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded  
 CC by it, and antibodies (Ab) specific for the proteins are useful for  
 CC detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body  
 CC fluids or tissue samples. HHV8 infections can be treated with antisease  
 CC or triplex forming molecules or agents that bind specifically to the  
 CC protein. Ab may be used for prophylaxis or treatment of HHV8 infection,  
 CC while the protein can be used in protective vaccines. Ab may also be used  
 CC to differentiate between lymphomas, and HHV8 may be implicated in many  
 CC other lymphoproliferative diseases such as lymphomas, leukaemia,  
 CC splenomegaly and mycosis fungoides. Cells and animals containing the  
 CC nucleic acid are useful for drug screening. HHV8-derived peptides can be  
 CC used as targets for antiviral drugs, e.g. dihydrofolate reductase gene  
 CC can be inhibited with methotrexate. These can also be used to determine  
 CC the immune status of a patient infected with HIV. HHV8 derived protein  
 CC viral MIP III may be used as an anti-inflammatory agent for, e.g.  
 CC treating rheumatoid arthritis. This sequence is stated as containing 81  
 CC open reading frames. (Updated on 27-AUG-2003 to correct OS field.)  
 XX

SQ Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 U; 0 Other;

Query Match

Best Local Similarity 17.7%; Score 1607.5; DB 2; Length 137507;

Matches 4481; Conservative 0; Mismatches 3869; Indels 342; Gaps 21;

Qy 1 ATGAGAGTTCCGTGGATGGACCAACACGTCGACCAAAATCCGGGTTTCATCTTGAAAC 60  
 Db 64892 ATGACAGCTCGGAGGGTGCATGATGACGAGCCGCGCCGGTTCATCTCAAC 64951  
 Qy 61 ATGACCTCGATGCCAAGTCCAGGGTTCGCGATCAGTCAGTCGCTGCAAAATATA 120  
 Db 64952 ATGACATCGATGCCAAGTCCAGGAGCGTAGTGGAAACAGATTGACAGGTTGTCAAAATATC 65011  
 Qy 121 ACTACAGCCCAACCGGAAATGGGTTTGGTACGACCTGGGCTTGGATCCGGCTGAAGACTCC 180  
 Db 65012 ACTACCTCGCCCGGAGATGGGCTGGTATGACCTAGAGTTCCGATCCACTGGAGACGAA 65071  
 Qy 181 GGGCGGTTCTTCCGCTTACCGTTTATCTAATACGGGAACGTCGTGGTGGGAAGAT 240  
 Db 65072 GGGCCCTTTCTCGCGTTTCGGCATACGTATTAACGGGACCTGCGAGGACGGGNAAGC 65131  
 Qy 241 ACCAGCATATCGGCGCTGTACCAAAATTTAAATCGCTGTATCAGGGCGGACCAACCATTA 300  
 Db 65132 ACCAGCGTATCGGCGCTGTACCAATCTCAATCGCTTAATTAACGGGCGCTACAGTGGA 65191  
 Qy 301 GCGGCACAGCACTATCGGCTCGCTTAAGAGCTTCTGTCCACGATCTTCAGGCGCTTTT 360  
 Db 65192 GCGGCACAGCACTTTCAGGCGCTTAAAGTCTTACTGTCCACTATATATACAGGCTTC 65251  
 Qy 361 GGCTTTAAGCGCGACATCAATATAGCCGTGAGAAAGCTCATCAGACCGGACCGGTA 420  
 Db 65252 GGATTCAGAGCAGACACATTAATATCTGCCAGAGGAAAGTCCCAAGGTAACCTCAGTCC 65311  
 Qy 421 TCCATAGACAAATTCAGCAACAGGAGCTATCGAAGTATTGGCGGTTTATAGTGACATT 480  
 Db 65312 TCCATCGACCACTCCAGAGTACGAGCTGGCTAGGTACTGGCCAACTGTCCACGATATT 65371  
 Qy 481 ATGAAAGAGGTTATGCGGAAACCAACCCATGGCATGTTCGGGACTATATCCACGCGAAT 540  
 Db 65372 ATTCGAGAAATTTATGCGCAAGAAACAAAAGGGCGGAGTATAGTCCCTCTCTCAAGCGCT 65431

QY 541 TTGAAACCTCTCGAGAAATACCGGACCGTGTGTTATGAGACTTCCAAATATTTATGTAATC 600  
Db 65432 TTCAGACTCTCTTTCGGTATGAGTGGTGGAGCCAAATTTGTGGAGAGTAACATATTCGTGATA 65491  
QY 601 GACGAGCGCGAAACCTCTCTCTTCAATCTACCAACCGTGTGTTCTTTTACTGGTTC 660  
Db 65492 GACGAAGCTGAACCTCTCTCTCCATATTTTGAACCGCGGTGTGTTCTTCTATTTGGTTT 65551  
QY 661 CTAAACAGCTGGCTAAATACCCCTCTTACCGCCAGCGGGCGGTTCGGTGCATAGTATGC 720  
Db 65552 TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTCCGCTGCATAGTCTGC 65611  
QY 721 GTCGGTTCGCCAAGCGAGACAAACGCGTTCAGTCAACTTACCAACCGGAGCGCAAAAG 780  
Db 65612 GTGGGCTCTCCACCCAGACGAGCGCGCTTTCAGTCCGCTCTTCAACCAACCGAGCAGAGA 65671  
QY 781 ACGAGATATCGTGTGCGAGAACATCTCAACATTTATGATAGGAGAGAGTCTGATCT 840  
Db 65672 AACGAGATATCTGCTGTGATATGTGCTACCTTCTCTATTTGGAAACCGTAGGTTGCA 65731  
QY 841 GAGTACGTACACTTGGAGAGAAACGTGGCGCTGTATTAAACAAAGCGCTGCACCGAT 900  
Db 65732 GATTATATTAGGCTGGACGAAATGGGCCCTATTTATAAACAAATAGCGCTGTACGGAT 65791  
QY 901 CTGCAGTTCGGACACCTTCTAAABATTTTAGAGTATAATCTTCCCATTTCTGACGAGTC 960  
Db 65792 CCCAGTTGGTCACTTCTGACAGACCTTAGAATATATCTAGACATATCACAGAGTTA 65851  
QY 961 ATGAGTTACGTAGACAGATTGTGCTTCTTAAAGTAAGATTATGATCTTTTAAATATAC 1020  
Db 65852 ATGACTATATAGATAGAGTTTGTGTTTCCGAAGAGTAAGATTCTGACCGCTCGAGTAT 65911  
QY 1021 ATTGGCTGAGCCGACCTTCTTTGTGCATAGAGGTAAGCGTAAAGCGTATCTAACACCTA 1080  
Db 65912 GCAGGTTGGAACAGACTCTTCACTCAACAGAGGTGAAGTCTTTTCTGGACAGCTG 65971  
QY 1081 CACACATGTCTAAACGTAGGCGGCGATAACAGGGAACAGAGTCTTTTACTGTCCCGTG 1140  
Db 65972 CACACCTGCTGTGAGTAATAGGATGCTGTGTCCCAAAAGCTTTTCACTGCGCAGTG 66031  
QY 1141 GTGTGCGAGGTGTTGTGAAGCGTTTGAAGATAACAAGCGGCGCTCACTCAACCAAC 1200  
Db 66032 GTCTGTGAGGTGTTTACAGAGCAATTTGAGAGGTAACAAGCGGCTAGGCTTCAACAC 66091  
QY 1201 CTCACGTTACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 1260  
Db 66092 ATGACTCCCAAGAAATGGGTAAACAAAAATCTTTTCAGGCTAAGTAACTACTCGAGTTT 66151  
QY 1261 GTGACAGAGACATGTCCATAGTCGCCAGCGAATCCACTGAAGCTTTCGACCCAGGTTACC 1320  
Db 66152 GCTGATCAGACATGGCTGTGGTGGACCTATATCACAGACGCGTCCACAGATCACCC 66211  
QY 1321 TTTATCACCAGTTGTTTAAACAGCCACGTCGTCTTTTAAATGGAATAACAAAAAATGC 1380  
Db 66212 TTCGCCACTAAATTTGTCAAAAAACAGCTATGCTTACCTTACTGGAAGAGCCAAAAATGT 66271  
QY 1381 ATATGTGGTTTACAGGTAGCTACTTTTGTAGTTTCAABAGAACTCTAGACAGCGAATCTTTC 1440  
Db 66272 ATATGCGGTTTACAGGTATACCAAGAAATTAAGTCCATCTAGACGCGGAGCTATTT 66331  
QY 1441 GTGAGACGATTCGCAAGCCGTCGGAATACGTTATACGTTTCTTAAACACGCTACTG 1500  
Db 66332 ATCGAAAGTCATTGCAACGATAACCCCGCTTATGTGTACAGTTTCTTATAGTACCCCTGCTA 66391  
QY 1501 TACAACGCAATGTAATCTGTTTACGCTACGCGTACGCGTACGTTACACAGAGAAATATCTG 1560  
Db 66392 TATATGCCATGTAATCATTTTACGCGCACGCGGTGAACAGGCGGATGAAGAAATCTCTC 66451  
QY 1561 CAAGACCTAAATTTGACCCCTTCCCGGCGCTCTTGGCAACCGGCGGTGTAGACCTTCAA 1620  
Db 66452 AGGACCTCAGGAACTCGCGGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 66511

QY 1621 ACGGTTGTTGAGAGTTAAACCTTGGAGACGACATCTTTTACCACTGTGTGCTCCCGG 1680  
Db 66512 GTTCTGGGCGAGGAGGGGACACAGATG-----CCTTCTACTCTACCGCAGCTCCCA 66565  
QY 1681 CCACCCCGGGGTATCACTCCCTCCAGGTTTGTGTGACACGTACTGTGCGCCCTAAAGGAC 1740  
Db 66566 CCATCCCCCAACCCACGCGGTCTTCCAAACACTGTGTGCTTATTACTCCGCGGCGCAAGAA 66625  
QY 1741 GTGTTCCGCTCCAGAAATAAGGTGGGTGTGCTGTTTGGCGGGAGTTTGAAGAGGAA 1800  
Db 66626 CTATTTCTGCAACAGGCTGGCCCTGACAGCGCGACACTTGTGTGACAGTTCTTCCACTCC 66685  
QY 1801 ACGTTTCCGCGTTTACGTTTAAACATGTGTAAAGGGAAGGAGTTGACTTTGTCTCCCT 1860  
Db 66886 GATTTTCAACAGTTTACGTTGAACATGTGTGCGAGATGGCGTGACTTTGTGTCCAAT 66745  
QY 1861 TCAGAACGTCTCAACGCGCTGTGGGTGTGCTGATCGACCGTTCGAATCGATATAAATAAG 1920  
Db 66746 TCCCCCGGCTCCACGCTGTAGTGGGATACGCACTCACTATAGACACCTATATAATCCAG 66805  
QY 1921 GGGTACAGTTTTCACGTTAGGTTGCGTGTGCTGAGGGTTTGCACCTCACTAGTATGAC 1980  
Db 66806 GGATATACGTTCTCCAGTGAGATTTCGCGCGTCCAGAGGAGCAGCGCCTCAGCAGGAC 66865  
QY 1981 CTCAGGAAGAGATGCGCTCCCTGCTGTCAGAGACTCTTAGCGGTTTATCGGCTCCCTA 2040  
Db 66866 CTGCGCAGAAAGATGCGCTCCATGTGTGTCAGAGTCTATCGGGTTCTATGCTGCTG 66925  
QY 2041 GAGATTAACATAACCAATTTGACCGAAACATGAGGAGCGGAGCGTTTTCOAAGTGTGC 2100  
Db 66926 GAAATAACCTCAACCAAGATGACAGAGACCTTCGAAGGTGGGACCGTGTTTAACATATGT 66985  
QY 2101 TGTGCGGGGACTATGCGGTTCAGTCAAAATTTAGCCATGACCATCGTAAGGACACAGGCA 2160  
Db 66986 TGTGCGAGGAGCTACGGTATCAGTTCTTAATCTGCTATGACCATAGTAGGAGCAGGAG 67045  
QY 2161 ATGTGTTGAGGAGGAGTAGCGGTAGTATTTGGTCCCAAGAAAGTCCAGACAGCCAC 2220  
Db 67046 GTTTCATAAGTAGGGTGGCCATATGTTTCGGCAACACCGCAATATCAGAGCCAGTCTA 67105  
QY 2221 GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTTGTGTGATGACAGACACCC 2280  
Db 67106 GTGTATGTGGGTGTATCCAGGGCCATCGAGCTCGTTACTGTGTATGACAGTAAATCC 67165  
QY 2281 CTTAAACCCCTCTCTAGAGAAACAGTTCGATACACCTCCGCAAGATATATGTCGCGCC 2340  
Db 67166 CTTAAGCTAATGGACCGCGGTGACGCCAGTCCCATCTCTCAAGTACATCATCAAGGCC 67225  
QY 2341 CTCACAAACCCAAACACACCTCTACTATA----- 2373  
Db 67226 CTATGCAACCCCAAGACTCTCTGATCTACTGACCGGTACCCCTCTCTTAGGACACTGAT 67285  
QY 2374 -----ATAATAAAAAACAAGAAATGGAATAAATGTTG 2406  
Db 67286 GTGTTTGGGAATTAAGCATGAGACTTGCACCTATAATGGTCTGTATTGACACCATTTCT 67345  
QY 2407 TTTTATTAGTCCAAACACCGCCACCGGATAGTTGTCAATTTCCACACACCGGGGGTGG 2466  
Db 67346 TTTATTATAGTCCAGCCACCGCGGATTTATGACCGTTTCCACAGCGGTTGCGTGG 67405  
QY 2467 CCGCCATAGTTTGAACACAGAAACCCCGGCCCTCTGGGTGTAGAGTAGGCC----- 2521  
Db 67406 AGGCCAGGATGCGGGTGGGTGCGTCACTCGACCTCGACCCCGCGGTAGTTGTCTCTGATG 67465  
QY 2522 -----GCCCGCGGTGCGAGGTCAATACGTCCTCTTAGGGCCCTGGGTT 2564  
Db 67466 AATCAGATGGCGGGAAGTACTGGGAGATTGGTTGGGAGGTGACCTTTGTGCTCGAG 67525  
QY 2565 GTACCAACATAATACAGTCTGATCATCACCGTTGCAAGTCCCTCGGTGACGCGGCTCTTG 2624  
Db 67526 GAGACACGATCAGCTCAACGCGGAGCGAGGCTCTCTGCTGTGCTCACTCCCGGAGATA 67585  
QY 2625 CAGGCGGACGACACCGCGGCTTGTAAAGTATCTG-----CAT 2664





Qy	4645	GGAAATATAAGTCACCTAAGTGGTATATATTTTCAACATGGGATTTACTACAAGCCCGGTA	4704
Db	69803	GGCAGGGTAGAGGCTCCCGATGCCTGAAATTGGCGCAACATGGTAAAGCGCACGCTATGCGTGA	69862
Qy	4705	CAACATGTATTAATAAABAAATCCACATGTAAATAATATAACTTAAGGTATACAT---	4761
Db	69863	GATGTGACCAATAGGGTGGTCCACAGGACGGCAATAGCGCAAGATCCCACATGGGGCAA	69922
Qy	4762	CTCATGTTTTATACATAATGAACGCTAATCGTACTTATATATATATATATAGAAATTA	4821
Db	69923	ATCCGGGTTTCACCCTTGTTGGCTGGTTCGGTCTCCAGGAGGCCCTTCCGTA	69982
Qy	4822	TGCCAATACAGTCATCGTAAAGTATATTTTAAAGGCATATGCAAAACATCAATATATAAC	4881
Db	69983	TATCTGTTTTATATAGTAGGGGTTACGCATCGCGAGTCCCGACTAATAGGCAATTA	70042
Qy	4882	ACAAAATGCAAGCTTCCGGATATCCGCCAGCTGTAAACCAGAAATACAGAAATGACTA	4941
Db	70043	CTGAAATGACCTTTTCGGCACACGGGGGTGAGGTCTATTTCCACGACATCTTCGCG	70102
Qy	4942	CACAAACACACCTGAAACCAATTTTATCTCAA-----CATATGCAACAATATT	4990
Db	70103	GAAAAATAGCCAGCTCTCTTAATTTCCGTGGGAAGACGATGGGGAAATGTGGCATTACC	70162
Qy	4991	TCAGGTAACACCATGTTAATAAATACGACATCGCAATTTTATAGTATAGCTCCTCAA	5050
Db	70163	TGACACGGTTTCAATCATACTCATCGTCGGAGCTGTACACGCTCGGCTGAGATTTTCTA	70222
Qy	5051	AGACAAATTTCTCTTCAATTGATGACTCATCAAC-----CTCGGTTTCCGTGTTAAATGTC	5105
Db	70223	AAAAGTCATCCCATGAATCATCGGAATCATACGACACTCTAGAACTACTCCATATGCG	70282
Qy	5106	GGAAATATATGATGTCAGAGTA-CCCTCGTCAATCGGCACAAATTAATTTCCAGGGCTTGA	5164
Db	70283	GGGTGGCGGGGTCGCCAGTAGTGACGCTGCCCATCGGGAGACACAGATGATGGGTTG	70342
Qy	5165	AAATATCATCCAGTGAATCAGCTGCTAAATGTCAAGGATGTTGTTCTAATTTTAGCCCAT	5224
Db	70343	AAATGTCATA CGGGCGGTGTGCAAAAGGTCACGTCCCATCCCAACACAGAGACCTT	70402
Qy	5225	CCAAATGGTAGTGGGTGACACTTTGGGCAAGAACCCTGAATAATTTGCAACGCTCCAAC	5284
Db	70403	TAGATACCTCTCCCGCATGTGCGGATTCGGGCAAGCAGCTGGTGTCTCGATATCC	70462
Qy	5285	AAACGAGGATGCAATTGATCGTTACAGGGGCGACATATTTATTTTATAACATCAG	5344
Db	70463	AAACGTGCCAGCGGTACCCAAAATCGCCAGGGCGTGTATTTATTTTCCACAGGAACCG	70522
Qy	5345	GAACCGGACCGGTCAATAGCACATATGTCATTAAAGATTGGAGTCGTGACACTGGTTG	5404
Db	70523	GTTTCTTAATTGCATCACCGGGTATCCAAAGCGGGCTTCCAGTTGATCCGGCTTA	70582
Qy	5405	AATTGACGGTGAAAAAATGTATTCGTGTACATCGGTAACATATTTTTTTAACACACGA	5464
Db	70583	CCGACAGTTCTTTTCCAGGGTTTCCGTGGGGCGCGGACGTGACTCAAAAAGGTCCTG	70642
Qy	5465	GGCGGAAATAAACGGGTCCCGCGCTAAGGCGGTGAATCGGCAGGACCGGCTCATTG	5524
Db	70643	CCTCTGCCCATGGCGGGTGGGTACAGTCCGCCATACTCTTCCAGGACACTGGCCATGC	70702
Qy	5525	CCATTTCCAGATTTGTTAAATCTGTTTCGAACATACGAGTCTAAAAATCAAGTATAACACG	5584
Db	70703	ATGACTCCAACCGTCTCACGTCCAGGTAATGTGCTCTATGAAGATGTGGTAGAGCCAGC	70762
Qy	5585	ACAAGTTCAACAGGAGGCAACGGGCTGGTTTCCACCGGACATCCCGCTCTCAAAACC	5644
Db	70763	AGAGGTTCAACACGATGAAATCAAGGTAAGTCTCCCGCGGAATCTCAATCCACAAGG	70822
Qy	5645	CGAACTCTCTAGGAACGCAATCAAAAGCGTTACAGACAAAATTTCCGAAACCGAAACCTTAG	5704
Db	70823	GGTATTGCTCCGGTGTCTGTATTAGTGCTTGGAATAGAAACTCAGAAAAAGACACTGACC	70882

QY	5705	TTCTGAGCACCATCTCTGGTTATTAATACTGGCGGTTAAGACAGACGACGCGCCACACACTGCTGCT	5764
DB	70883	CACCAAGGAGAACTTGGCGCTCTTGCAAGATGTGATAGACGCCGCGCAAGAAATGTTCTTCC	70942
QY	5765	TGTGAGGATTCAAGCTTAGGTGCGCTTGCAGGCTTGTCTGACAGGCCCGTAGCTGCCCTGGCGG	5824
DB	70943	CGTGGACAAAGAGCTTGGGGGGGACAGATGGCCCTACAGTGGGTGATTTCTTCTACCA	71002
QY	5825	CGCTCGCGCACCTTTTACAGCGCGACGCGAGTTTCAGAGCGAGGACAGAGCGCTCTTTTGC	5884
DB	71003	CGGTCTATACATTGGTGGCACCCACAGGCGCTGTTCAGTATCAGCAATAAATCTATCTTTTGC	71062
QY	5885	AGTCGGCGCATCTCTCGGCATAGAGGATTCTCGCTCGCTGATCTTTTAGGTGTCATATATA	5944
DB	71063	AGTCATCCAGATCAAAAGTCATGTCCAGATGCTGTGGCTGGCAATTTGCCCGCATGTACA	71122
QY	5945	TCTGTTGGGTAAACAAACCTCACTCCCTGTAAAGGGGAATAAGTCCGTTCTTTCTATCG	6004
DB	71123	TTTCCGTGCCACATATTTTAACTCTGTAACTGGAAGTAGATTCACTCTGGTGTGA	71182
QY	6005	TTTTCTCTGAATCTATAGTTACATCTTTAAGTATTAACCAAGACGCTTAAACCCAGAT	6064
DB	71183	GCCTCCCGGGGAAGCCAGCGTATGCTTCAGACACCAAGGAGCGCTAAGAACCCCGGT	71242
QY	6065	GGTCGACACCTTTTAAAAAATCGAAAGCGACCGATGGCTCTAATCCGCGTTTCCATGTC	6124
DB	71243	GTCCGCGCTCCGAAACAGACCTCTGAGATAAGTCTGGTCTTGACGAAACCCGATGG	71302
QY	6125	GTTCAAAGGTAGTCAGATGATCTCCAGTTAGACTGTGTCTCCCTGTTTAAACCCCTTAA	6184
DB	71303	TACCGAATGCCAATCTCTGGCCCTCCAGCTCTCAAAATTTTCACTCCCAATACCCGAA	71362
QY	6185	CTGGAATGGAACAGACCATGACACCGTGAACCTTCTGATGCTCTCTAAACTTAAACCCGA	6244
DB	71363	TTGGGATACHACCTCCATGTTCACTCAGTCACATGTAGCTAGGTCTCCCCACCCACCCC	71422
QY	6245	AACAGAGCTAAATPACCAATGACTGTCAACCCCTACCACCAAGCACGCCCCCGTACTATTA	6303
DB	71423	ATAGGACCAAGTACAGCTTATCTCTCAATAATACCAGGAGCTACCGGGAGCTCATTA	71482
QY	6304	-----GACCGGTGAGTAAACCGCTATCCCTTAAATAACCCATAC	6344
DB	71483	AGCCCCGCCAGAAACCAAGTACGTGGGTGGCAATGACAGTCCCTTTTAAAGATCAACC	71542
QY	6345	GTGGAGTTTGTAAGGTAAACAGCT-CGTGTATTTTACAGCGCTGTGCATAAATGGATA	6403
DB	71543	TTACTCCGCAAGGGTAGTCTGTGTGAGAACTACTGTCCAGSAGCCACAAAATGGCG	71602
QY	6404	CCGACGCAATCAGGTAAATAACTTTTATTCAAGATTCAGAGGCGGTGTTTACAG	6463
DB	71503	AAGATGCAAGGTAAAGATCGACC--TTTTTATTGTATCTAGTAAACAATGCGGTGTTTCAA	71660
QY	6464	TGTTGATGTTGGAGCATATCTCGCAATGGAAATGGCTGTATGAATCCACACTTAGTGCT	6523
DB	71561	TGTTGATGTTGGAGCAGAGTTCGCCAGCTCTACGTCCGAACAGTCGGGTGTCAAGSCT	71720
QY	6524	CGGATCAGGTGCTCTGTGTACTTTTACTAGTCCGATGAACCTTAAATTTGGCGGTGTTAGG	6583
DB	71721	CTTTATTAAAGTTTCGGTGTACTTGAACCAAGCGCGGAACCTTAGGTGGGTCTGTACAGG	71780
QY	6584	TCGTACCAACCAAAAATCTGCTGGCTCGCTTTTAGAGCTCTATCGAAGCTTTTAAATC	6643
DB	71781	TCGTACCAAGCAAAAAGATTCGGGGGTGCTTTTCAGAGAGTTAGGAACGTGCTGAT	71840
QY	6644	GTGTCCAAATAGCTTGTCTTTAAAAATGCAACTCTGGTATCATCTTTATGACAGTGTCCAA	6703
DB	71841	ATGTGGACAGCTTCTGCTCGTAAATGCAACCGCTGGTATCTGAACAGACAGCTGTCCAA	71900
QY	6704	AAAAACAAGATTTAAAAACAGTTAAATCTCGGTGTCTCTGTAGTCTCTCGTATATAACC	6763
DB	71901	AAAAACAAGGTTCACTGTCACGTTAAATCTGTATCTCTGAAGTCTCTCGTAAATGACA	71960
QY	6764	TCCTCAACAAGAAAAATTTTTTAAACCAACTCGCCAGGTACTGAAACGATGGTACGGAC	6823



Db	71961	GTCTTACCAAGAAACCTTTTACCACTGCGCCATCCACTGAAAGGAGGACACAC	72020	73041	TCTGCAATCGAGGCCCATTAATGATATTAACCTCTCCCTACTGAAGCTCTACAAACGCT	73100
Qy	6824	AGGTCGCGGAAAGTGTCTATCATCTCTTCACTCCCTCCCTTTGAAGGTTTGGTTACACCG	6883	7893	TGGATGCCGAGGAAAACTTGAACCTTAGACTCAAGACACTGTGCAAACTTTGTTTCAAC	7952
Db	72021	GTCCGTTGCGGTTGTAGGATATCCCTAACTTCGAGCGGAGACGGCCGACGCTCC	72080	73101	ACGATACCGCGCGGCTGCTCTCTGAGCAGTCCAGGGCCCTCTGCTTTTGGTTTCCAAAC	73160
Qy	6884	ACAMTGTGAGATAGGACCAAGCTCAGGGGGAGAGTCCGATCGTGAAGGGGTAAACATTCG	6943	7953	CCATCTTTCTTCAAGATCTTAGCCCTATGCACTACTGGCCCAACGGGAAGATGCTC	8012
Db	72081	ACAAAATGGAGAGGACCACTCTGTGSCAGTCCCGGTCTGGGGTTCTGATTCACAGGGC	72140	73161	CGGTCTACCTTCCGAGGATTAAGCGCGCTGGAGATCATGACCAAGGTCAGCTCGCC	73220
Qy	6944	TGGCGGTGGGGTACTCCAGTTTATAGATCTCCGGCAGGACGCTACGAGTTCTTCTGTC	7003	8013	CTGCAATATTTTACCATCACCGGTACGGCCGAGAGAGCGCCCTGTCTGTAAGTGA	8072
Db	72141	GCGGTGTGGGGTATTTGAGAGTCAAACTCTCGGCAGTCCCTTAATGAGCTCTCTCA	72200	73221	CTGAAAACCTTTTACAGCATCACCGTTCTGCTGAGAAAACGCGGCCAATACCAACG	73280
Qy	7004	AATCTACGAGCGCCGATCCACAGGGGAGTACTGAGTTAATAAGTCCCGCAATCTA	7063	8073	GTACTAGCGGATGACGTGTCCGGAAGCAGCTTGTCCCGATTCCTTATCTGCTCCAG	8132
Db	72201	AAACCTATGAGCGCGCTCACTAGTGGCAGCATGCCGTTAATAACACCCCTTATCTTG	72260	73281	AGGTCACTGGAAGTGTCTTATCCGGAAGCGGTCTCATGCGCAGAAATCTTTAATTTGCC	73340
Qy	7064	TCGTTCTGTAATTTGTGGAGCTCTTTAGACAGTAACCATGTTGCTCCGACACTCGCGT	7123	8133	TATGCGAACCGGGCTTCTCCCGGCACCCCTGGTGTGACCTCAGTAATGTCTTAGAAAATC	8192
Db	72261	TCGTTGCCAAGTTTGTAACTGCTGAGGGAATAAGCCAAATTCGCCCTAGCGCGGGA	72320	73341	TCCTGGAGCCAGAGCTGTTGCGGCTTCCATGGTAGACCTCAGCGATGTCTGCAAAAAC	73400
Qy	7124	ACGATGAACCGCTCTCTGTTGTTGATGAGCACCTGGAGTATCTTTAGACAGGTAT	7183	8193	CAGAAATCATCTCAGCGCCCAACCCCTGAGTCAATTTGTATCATCACAACAGCACCCCA	8252
Db	72321	ACCAGGTACGGCTCGCTTGTGCGTCTGGACCAATATCTGAATGCTTTTCAAGGTAT	72380	73401	CGCGCTTATTTCTGAGCGCCCTGCTGAGCCAGTTTGTCTATTAGCAAAACCCATCCCA	73460
Qy	7184	AGGCGCGTTCCACTCTAAGGTTGCGCTTGTGAGAGCGGATTCGGGCAATGACGGTTC	7243	8253	GTCTGCGCTCACTGATCAGCATTTATTAGCCCAACCCAGGGGCTTGTTCCTCCGCGCAATGT	8312
Db	72381	AGGCTTCTTCAAGGTTTGAAGCGGTACGTCGACGCTCTGGATGAGGGTGGCGACGAC	72440	73461	ACATGCCGCACACCGTCAACATCATCCCTTTAACC---CATCGGTACAGACCCCGGT	73517
Qy	7244	AGTGTGTTCAATTCACCGACTATCATGCCAGTGGGGATAGTGGAACTTTGTACATATGA	7303	8313	TTATGGACACGTGGAAAGCGGTGTACAGAGCATTACCCAGGCGACAGAC	8364
Db	72441	AGGTAATCTAATCTGAGTATCTGATCCAGGACGGGTAAATGATACCTAAACAGATGG	72500	73518	TTATTAGTACGTGGCAGCGCGGTACAGAAATATGGTGTACAAACATCCAC	73569
Qy	7304	TTTAAAGATGGTCTTCCCGTAATGGCGACTCATGTCGTGTTTAA-----AC	7352	RESULT 4		
Db	72501	TTGAACAGGTGATCTTTAAGGGGCTTCTCGATGTCATTTGAAAAAATATGACACGCCAC	72560	AAV73803		
Qy	7353	GTTCACAGGTTTAAABAACTCAGAGGTCCTGCGTGACAACTTTTGGGACTCTCG	7412	ID	AAV73803 standard; DNA; 35100 BP.	
Db	72561	TCTCTCTTAGGTAAGAGCTTCGGCGTCTCTGTGGGAACTTCTGCGGCTCTCG	72620	XX	AC	AAV73803;
Qy	7413	AACCGCTCAGGATAGCTTACCAACCGGTCTCTTTATAACGAAGCATACGTCAAAA	7472	DT	17-OCT-2003 (revised)	
Db	72621	ACGAACGAAAGGCCAACTTACCAGTGTGTCTCTTATAAATGACGCATACCAACAA	72680	DT	25-FEB-1999 (first entry)	
Qy	7473	AACATGACCCAGTACCTAAATAGCTCAGAGAGTCTCTGATGTATTGTATCG	7532	DE	KSHV LUR DNA (nucleotides 35,101-70,200).	
Db	72681	TCTACGATCCCGTACCTAAATAGATGTGTGAAGATGTGATCGGATTTATGAAG	72740	XX	Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;	
Qy	7533	AATGTGCTCCTGGTCCCATCTCCGGGCTCATCGCGATCTGAATCTCTTTAACCTGT	7592	XX	dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;	
Db	72741	AATGTTCAAGCTTGGTCGCTATGTGTTTACAGACATTAACATGTTTAACTTT	72800	XX	diagnosis; treatment; HHV8; thymidine kinase; glycoprotein H;	
Qy	7593	TCTGCTCTACCGGGTTCCCGGTTAAACCCCGGGGCGGCACCTGTAAGTCCCGT	7652	XX	capsid protein; packaging protein; helicase primase; glycoprotein M;	
Db	72801	TCTGCTTTTTCGTGCTCTCGAATGAGGACCAAGCGCGCGGTACACAGTGCCTAT	72860	XX	viral protein kinase; alkaline exonuclease; virion assembly protein;	
Qy	7653	GCGCAGATGCGCCAGCGCATCTGAGAAATCTCACCGAGAGCGCCCTTGTGTACCG	7712	OS	uracil DNA glycosylase; UDG; glycoprotein L; ds.	
Db	72861	GCGCAGAGGCAATCCAAAGCATTTTCGATCTCTCACCGAGAGATCTTATCTGCACAG	72920	XX	Human herpesvirus 8.	
Qy	7713	AAAAGATGTTATAGCTGTGCTCGAGCGGGTGTATACCGCCCGAGTGGCCAGAG	7772	PN	US5849564-A.	
Db	72921	AAAAGGCAATTTCTGACAGCGCATGACAGCGGGTGAAGCTTCCAGCCATATGTAAGC	72980	XX	15-DEC-1998.	
Qy	7773	TTCTGACGAGTCTACCGGAATGAAGCCAGTGTGTTGGGGGTGGCGTCTCA	7832	PD	29-NOV-1996; 96US-00770379.	
Db	72981	TACTACGAAATATACACTGAATGAAGGCCAAATGCTGGGGCTGAGCGGACTCG	73040	PF	29-NOV-1996; 96US-00770379.	
Qy	7833	TATGTTGAGCGGCGCTATATGCTATTGCCGACTCGGTCTCTGCTCACTTATATACCC	7892	PR	(UYCO ) UNIV COLUMBIA NEW YORK.	
				XX	Bohenzky RA, Moore PS, Russo JU, Chang Y, Edelman IS;	
				XX	WPI; 1999-069741/06.	
				XX	Kaposi's sarcoma-associated herpes virus nucleic acid - encodes	
				PT	di:hydro:folate reductase and is useful for treatment, prophylaxis or	
				PT	diagnosis of Kaposi's sarcoma.	
				XX		

PS Disclosure; Col 97-126; 109pp; English.

CC This sequence is a fragment of the Kaposi's sarcoma-associated

CC herpesvirus (KSHV) LUR (long unique region). This fragment contains

CC coding regions for ORF20, ORF21 which encodes thymidine kinase, ORF22

CC which encodes glycoprotein H, ORF23, ORF24, ORF25 which encodes the major

CC capsid protein, ORF26 which encodes capsid protein II, ORF27, ORF28,

CC ORF29b which encodes packaging protein I, ORF30, ORF31, ORF32, ORF33,

CC ORF29a which encodes packaging protein II, ORF34, ORF35, ORF36 which

CC encodes viral protein kinase, ORF37 which encodes alkaline exonuclease,

CC ORF38, ORF39 which encodes glycoprotein M, ORF40 which encodes helicase,

CC primase subunit I, ORF41 which encodes helicase primase subunit II,

CC ORF42, ORF43 which encodes capsid protein III, ORF44 which encodes

CC helicase primase subunit III, ORF45 which encodes the virion assembly

CC protein, ORF46 which encodes uracil DNA glycosylase (UDG), ORF47 which

CC encodes glycoprotein L. KSHV is a new human Herpesvirus (HHV8) believed

CC to cause Kaposi's sarcoma (KS) which is the most common form of neoplasm

CC occurring in persons with acquired immune deficiency syndrome (AIDS). The

CC DHFR protein is useful for vaccination, prophylaxis, diagnosis and

CC treatment of a subject with Kaposi's sarcoma and for detecting expression

CC of a DNA virus associated with Kaposi's sarcoma in a cell. (Updated on 17

CC -Oct-2003 to standardise OS field)

XX SQ Sequence 35100 BP; 8016 A; 9231 C; 9477 G; 8376 T; 0 U; 0 Other;

Query Match 11.0%; Score 1002; DB 2; Length 35100;

Best Local Similarity 54.5%; Pred. No. 5.4e-287;

Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;

QY 1 ATGGAGAGTTCGTCGGATGGAACCAACACGTCGAAACCAATCCGGGTTTCATCTGAAAC 60

DB 29792 ATGGACAGCTCGAAAGGTGCACTGATATGAGCAGAGCCAGCCCGGGTTTCATCTCAAC 29851

QY 61 ATGAGCTCCGATGCCAAAGTCAGGGGTGTCGGTGAATCAGTCAGTCGCTGTCAAATATA 120

DB 29852 ATGACATCGATGCCAAAGTCAGGAGCGTAGTGGAAACAGATGACAGGTTGTCAAATATC 29911

QY 121 ACTACAGCCCAACCGAAATGGTGGTACGACCTGGCTTCGATCCGGCTGAGACATCC 180

DB 29912 ACTACCTCCCCCGAGATGGCTGGTATGACCTAGACATCCATCCACTCGAAGACGAA 29971

QY 181 GGGCCGTTCTCGCTTTACCGTTTATCTAATACGGGAACCTGCTGGTGGGAAAGT 240

DB 29972 GGGCCCTTTCTCGCTTTTCGGCATACGTAATAACGGGACATGACGAGAGCGGGGAAAGC 30031

QY 241 ACCAGCATATCGCCCTCTACCAAAATTTAAACTGCCTGATCAGCGGGCGACCAACATA 300

DB 30032 ACCAGCATATCGCCCTCTACCAAAATTTAAACTGCCTGATCAGCGGGCGACAGTGTA 30091

QY 301 GCGGCAAGAACTTATCGGTCGCTTAAGAGCTTCTGTCCACAGATCTTCAGCGCTTTT 360

DB 30092 GCGGCAAGAACTTATCGGTCGCTTAAGAGCTTCTGTCCACAGATCTTCAGCGCTTTT 30151

QY 361 GGCTTTAAGAGCGACACATCAATATAGCCGTGAGAAAGCTCATCAGACCGGAGCCGTA 420

DB 30152 GGATTCAGAGCGACACATTAATATCTCCAGAGAAAGTGCCTCAAGTAACTCAGTCC 30211

QY 421 TCCATAGAGCAATTCAGCAACAGGAGCTATCGAAGTATGCGCGGTATAGTGAGCAATT 480

DB 30212 TCCATCGAGCAATTCAGCAACAGGAGCTATCGAAGTATGCGCGGTATAGTGAGCAATT 30271

QY 481 ATGAAGAGGTTATGCGAAACCAATGCGCATGTACGGGACTATATCCAAACCGAAT 540

DB 30272 ATTCGAGAAATTTATGCGAGAAACCAATGCGCATGTACGGGACTATATCCAAACCGAAT 30331

QY 541 TTTTGAACCTCTTCGAGAAATGACCGGACCGTGTATGAGCTTCCAAATATTTGTAATC 600

DB 30332 TTCAGACTCTTTGGCGTATGGGTGGAGCCAAATTTGTGACGAGTAACTATTCGTGATA 30391

QY 601 GAGAGCGCGGACCTCTCTTACATACATCACCACCGTCTGTCTTTTACTGGTTC 660

DB 30392 GACGAAGCTGGAACCTCTCTGTCCTCATATTTTACGCGCGGTGTGTCTTCTATTGGTTC 30451

QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTTACCGCAGGGGGCGGTTCCGTGCATAGTATGC 720

DB 30452 TACAACAGTGTGGCTGACACCCCGCTATACAGAAATGGTGGCTTGCATAGTCTCC 30511

QY 721 GTGGTTCGCCAACGCGAGCAAAACGGTTCAGTCACTTAACAACACGGGACGCAAG 780

DB 30512 GTGGGTCTCCCAACCCAGACGCGCTTTCAGTCCGTCTTCAACACACGCGACGAGAG 30571

QY 781 ACGGAGATATCGTCTGGCGAGAACATCTTAAACATTCATGATAGGGAAGAAGTCTGATCT 840

DB 30572 AAGGAGATATCTGCTGCTGATATATGCTCATCTTCTTATGGGAAAACGTTGAGTTGCA 30631

QY 841 GAGTACGTACACTTGGAGAGAACTCGGCGCTGTTTATATAACAACAAAGCGCTGACCGAT 900

DB 30632 GATTATATTAGGCTGGACGAGAAATTTGGGCCCTATTATTATAACAATAAGCGCTGTACGGAT 30691

QY 901 CTGCACTTCGGACACCTTCTTAAATTTTAGAGTATAATCTTCCCATTTCTCTGACGAAGTC 960

DB 30692 CCCAGTGTGGTCACTTGTCTGAAGACCTTTAGAAATATATCTAGACATATCACCAAGATTA 30751

QY 961 ATGAGTATACGTAGACAGATTTGTCTGTTCTCTTAAAGTAAGATTATGATCTCTTTAGAAATC 1020

DB 30752 ATGCACTATATAGTAGGTTTGTGTTCCGAGAGTAAGATTCTGACCCGCTCGAGTAT 30811

QY 1021 ATTGGCTGACCCGACTCTTTTGTTCACATAGGAGGTAAAGCGTATCTTAACAAACCTA 1080

DB 30812 GCAGGGTGGACAGACTCTTCTATCTCACACAGGAGGTGAAGTCTTTTCTGGACAGCTG 30871

QY 1081 CACACATGCTACAGCTAGGGGCGGATACCCAGGACACGAAGCTCTTTTACTCTCTCCCGTG 1140

DB 30872 CACACCTGCTGTCTGAGTAAATAGGATGCTGTGTCACAAAGCTTTTCACTCCCGCAGTG 30931

QY 1141 GTGTGCGAGGTGTTTGTGAAGCGTTTGAAGAAATAACGCGGCGCTCAACCTCACCAC 1200

DB 30932 GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGGGCGGTAGGCGCTCACAC 30991

QY 1201 CTGACCGTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTTTCGCAAGTTT 1260

DB 30992 ATGACTCCCATAGAAATGGGTAAACAAAATCTTTTCAAGCTAAGTAACTACTCGAGTTT 31051

QY 1261 GTGGACACGACATGTCATAGTCGCCACGGAATCCACTGAAACGTTGCGACCCAGGTTACC 1320

DB 31052 GCTGATCAGACATGGCTGTGGTTGGACCTATATCACAGACGCTCCACACAGATCACC 31111

QY 1321 TTTATACCAAGTTTGTAAACACGCCAGCTGCTTTAATATGMAAAACAAAAAATGC 1380

DB 31112 TTGCGCACTAAATTTGCAAAAACAGCTATGTAACCTTACTGGAAGACCAAAAATGT 31171

QY 1381 ATATGCGGTTTCAAGGTACGTACTTTAGTTTCAAAAGAAATCTTAGACGCGAACTCTTC 1440

DB 31172 ATATGCGGTTTCAAGGTATACCAAGATTCAGTCCATCTTAGACGCGGAGCTATTT 31231

QY 1441 GTGGAGACGATTCGCAAGACCGTCGGAATACGTATACGGTTTCTTAACACGCTACTG 1500

DB 31232 ATCGAAAGTCAATTCGCAAGTAAACCCCGCTTATGTGTACAGTTTCTTGTAGTCCCTGTA 31291

QY 1501 TACAACCCCATGACTCTGTTTCAAGGTACGCGGTGACTAGGTACACAGAAATATCTG 1560

DB 31292 TATATGCCATGTAATCTATTTTACGCGCAAGGAGTGAAGCGGCGCATGAAGTATCTC 31351

QY 1561 CAAGACCTTAAATTTGCAACCCCTCCCGCGCTCTGCGAAACCGGGCGGTGAGACCTTCAA 1620

DB 31352 AGGACCTTCAGGAACTTCGCGGTGCTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGAC 31411

QY 1621 ACGGTTCTGAGAGTTAAACCTTGGAGAGACGACATCTTCTACACAGTGTGCGAGTCCCGCG 1680

DB 31412 GTTCTGGGAGAGAG-----GGGACACAGATGCTTCTTACTCACCGCAGCTCCCA 31465

QY 1681 CCACCCCGGGGTATCACTCCCTCCAGGTTTGGTGCACAGTACTGCGGCGCTTAAGAC 1740

DB 31466 CCATCCCGCCACCCACCGGGCTCTTCCAAACATGTTGGCTCTTACTCTCGGCGGCGCAAGAA 31525

QY 1741 GTGTTGCGCTCCAGAAATAAGGTGCGGTGTCGCTGTTTGGCGGGGAGTTTGAAGAAGAA 1800

31526 CTATTCGTCAACAGAGCTGGCCCTGGACAGCGGACACATTTGGTGACGAGTTCTCTCCACTCC 31585  
1801 AGTTTTCGCGCTTTACCGTTAAACATGTCCTTAAGGACGAGTTGACTTTGCTCCCT 1860  
31586 GATTTTCAACGTTTACCGTGAACTGCTGTGGAGATGCGGTGGACTTTGTGCCACT 31645  
1861 TCAGAACGCTCAACCGGCTGTGGCGTTTCATCGACCGTTGAATCGTATAAAATTAAG 1920  
31646 TCCCGCGGCTCCACGCTCTAGTGGCATACGATCCATATAGACACTATATAATCCAG 31705  
1921 GGGTACACGTTTACCGTACGCTTCCGTCGCTGTCAGGCTTTGGCCACTCAGTGATAC 1980  
31706 GATATATCGTTCTCCAGTGAGATTCGCGCTCAGAGAGACAGCCCTCAGGAGAC 31765  
1981 CTCAGAGAGAGATGCGCTCCCTGCTGTGTGAGAGCTTACGCGTTTATTCGGTGCCCTA 2040  
31766 CTGCGCAGAAAGATCCCTCCATAGTTGTCCAGGACTCATCGGGTTTCAATGCGCTGCTG 31825  
2041 GAGATTAACATAACCAAAATTCACGAAACCACTGAGGACGGGCGCTTTTCCAAAGTGTC 2100  
31826 GAAATTAACGTCACCAAGATGACAGAGACCTTCBAAGTGGCGAGTGTTTAAATATGT 31885  
2101 TGTGCGGGGACTATGCGGTGAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 2160  
31886 TGTGAGGGGACTACGCTATCAGTTCTAATCTGGCTATGACCATAGTGAAGGACACAGGG 31945  
2161 ATGTCGTTGGACGAGTAGCCGTAGTATTGGGTCACCAAGAACGCTCCAGACAGCCAC 2220  
31946 GTTTCACCTAAGTAGGCTGCGCATATCGTTTGGGCAACACCGCAATATCAGGACGCTCA 32005  
2221 GTGTATGAGCAATATCAAGGCTGTAACTCAAAATTTATGTCATGAGCAACACCC 2280  
32006 GTGTATGTTGGTGTATCCAGGGCCATCGACGCTCGTTACCTGTGTAATGAGCAATAATCC 32065  
2281 CTAAACCCCTCTCAGAGAACCACTCGATAAACCTCCGCAAGCATATAGTCCGCGC 2340  
32066 CTTAAGCTAATGGACCGGCTGACGCCGCTCCCTCCCTCAAGTACATCAAGGCC 32125  
2341 CTCACAAACCCAAACACAAACCTCATCTACTAA ----- 2373  
32126 CTATGCAACCCCAAGACTCTGATCTACTGACCCCTGACCCCTCTCTTAGGACACTGAT 32185  
2374 -----ATAAATAAACAAGCAAAATGATGCTGTATGACACCTCTT 32245  
32186 GTGTTTGGGATAAAGCATGAGACTTGACACTATATATGCTGTATGACACCTCTT 32245  
2407 TTTTATTCAGTCCCAACACCGCCACCGGATAGTTGTCAATTTCCACACACCGGGGGTGG 2466  
32246 TTAATTTATCAGTCCAGCCACCGCCAGTTATATGACCGCTTTCCACACAGGGGTGGCGTGG 32305  
2467 CGCCATAGTTTGACGACCAAGAACCGCGGCGCTCTGGGTTAGAGTAGCCCC --- 2521  
32306 AGGCCAGGATCGGGTGGTGGCTGACCTGGACCCCGGGTAGTTGTCTTCTGATG 32365  
2522 -----GCCCGCGGTGAGTGCATACGTCTCTTTAGGGCCCTGGTT 2564  
32366 AAATCAGTGGCGGAAGTACTGGGAGATTGGGTGGGAGGTGACCCCTTTGTCTCGAGC 32425  
2565 GTACCAACATAAATCAGCTGTCATACCGTTGAGTCCCTGCTGAGCGCGGCCCTCTTG 2624  
32426 GAGACACGATCAGCTCAGCGGACGAGGGCTCTCGTGTGTGCTACTCCCGAGGATA 32485  
2625 CAGCGGACGAAACCGCCCTGTGTAAACGTAATCT -----GCAT 2664  
32486 TAAATATCAGGACGCCACTGTTTGGGCTTAAAGTTTGGTTGTCTGTCGACGCGACCA 32545  
2665 GTCGTGCTGTGCTGAAGATCAACACAGGTCAATCACAATCAGTGGGCGCCGCTCCCTCAT 2724  
32546 CATCTCTCTACAGAGAGGCGGTAGACTGCTCTTTGGCTTCTGGCCCAAGTCCATGA 32605  
2725 CGGCGTCTGATGCTGCCATGTCTCAACAGAGCTCGGTGCTCACTTCTGCTGCTATCTG 2784

32606 GCCGATTCTCTGACTCAATACTTCCCTTGGTCTTCTCGTCTCTCTCGACGAGGTG 32665  
2785 AAAATAAATGCTCTTGTCTCTGAGTTGATCAGCAACAAAGAGTTCTTCTTCATCGG 2844  
32666 GCTGGTGGAAAAATGCGCGCGCTCGGTAAACGCGGCTCATTTGTTCACTCCGAGAGT 32725  
2845 AAGACCATCCGCCCTCTCCCGTGATCTAGATATATACCTGACACGCGGCGGAGGCC 2804  
32726 TGGAACTGTCTATCGTATCAGAGTCCGATGTCAAGTCCAGATCGCGGTGGGTGCGCGC 32785  
2905 G----- 2905  
32786 CAGGGGCGGCCACGAGGGGCTTTCATCAGGTCGCTGTATGTGTAACATTTGTGTTCCAG 32845  
2906 -----TGGGAAAAAGCGCTGAGCTTCTCGCCACTTTTCATCAAGTGT 2946  
32846 GTACACTATTTCTGGAAAGCAGGTGAAGTCCGTATGCCCGTCCCAAGTGTATGCCGCCA 32905  
2947 CTTGCTAGTCTCATATATAAAAAATTCGTCTTCTGCTGCTGATTCAGATGACTCGGGTGGC 3006  
32906 TCGGTTCCAGATAGCAACCCCTCTGCTCTGAAGGTGAGACCCAGACGAGGAAAAATC 32965  
3007 GGTGTTGGCGATTAACCAACCGCAACCGGC-----CGCGAACCAACAAACGGG 3053  
32966 CGTCATCTGACTAACCCATCCCATGAGACGCTCGGACTCCGCGCTGTCTCGTTGAACTGC 33025  
3054 GTCAAGTGGCGGCGCGCTCGGTGCTGTCGACCTGTGCAAGTAAATGACGAGTGT 3113  
33026 GCAAGCGGCGCGCTTACACATGCTACCGTTTGGGCGTATGGGCGCTCTGGCCAGAGGCC 33085  
3114 AGCGCGCGCTCTGTGTTAAAG-----ACTCAACCGTCCCGTG 3150  
33086 TCGGCGCAAGTGAAGTAAAGTTGAAAAAAGTCTGAGGTTACCCCTCTGGCTCTCTT 33145  
3151 CGTCGCGATCTATTTCTAGGTGGGAGGAGACCGCGGTTCTGCTCTTCTTCATCCACATCAC 3210  
33146 CTTCTCTGAAATCATCGTCAATTTCTTCTCATCTTCTTCATCTCTCTGCTATATTCAG 33205  
3211 TGTAGTCTGCTGTTCTCCGAATGAGAGTGGTTCGCTGTCGGCTTTTGTGAGTAGTGGAG 3270  
33206 ATTCCGCGCTCGACTGATCGGGGATATCTGTAGATCCAGAGGGTGTCTGGCGCGATG 33265  
3271 GGGAACTCTGGTCAAAAACATCTCAGGTGCTGATGTCATCAATCACTGGGGGCTCTCGT 3330  
33266 GCGTGTCTCTGGCGAGACGCTGCTGCGGCGACATATCTATCACCGTGGGTCCAGCAT 33325  
3331 ACCCGGCGGTACCGGCAACCCCGTACGGAA-----CCCGGCGAACAAAAAC 3378  
33326 AGCGCGCGGCTCGCCAAATCTCGAAGTGTGAAAGAGTGGAGTGGGAATATGAAT 33385  
3379 GTCTTGTCTCTTTTCTTAGGTGCGCGGAATCGGACATCTCTGGGAGTGGCGCGCG 3438  
33386 TCACGGGGGTGCTCTGCGAGGCGCTCTTCAATTCGAAGCATCTCTCTTCTCATCGTGTG 33445  
3439 GGGTCTTGTGACGGGTGGATAAAGAAC-----ATAGCATGGCC 3476  
33446 TGCTAGACGAGGTCTCTCAAAACATCGCCATGCGCTTTGTACGGGGTTGACCGTGGGCG 33505  
3477 GAAACGTTCACTGTATAAAGCAGCTGCTATCCCGCATCAGTTAATATTCAGTCAACGG 3536  
33506 GGAATTTACAAGACACACAGTTATTGCTTTTACTGCTCCACAGCGGCCAGTCCACAG 33565  
3337 CACCCCTCCGATTTCTGAATAGTATTCATTCGAAAGTTTAAATAGTTGCGGCTTAGAA 3596  
33566 TCTCACCGCGGTGGCGAGTCAAAATAGTCTGTGGCTAGGTAAAGTGAATACAGCCCTGA 33625  
3597 ACTCGCGCAGGAGAACCGGTTGCGCGCTGGCGGTAGCCAGGGGAGAGCGATGGCGCG 3656  
33626 ACCGAGCCATCGCGAGTGTGGCCACCAAGAGAG-----GCCAGCGAGATGATGCTGG 33682  
3657 CTTGAGGAGAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3716  
33683 CCGTAAGCACAGGTGTTCTGTGCGTTTATGAGCGGAGTCTCTCAATGGGCTTGGCGCC 33742



Qy 6231 AAAAATAACCCGGAACAGAGCTAAATACCAATGACTGTACCCCTACCCGAGCCAGC 6290  
 Db 1209 CCCACCCAAACCCATAGGACCCAGCTACAGCTTATCTCCCTAAATACCAAGGAGCTA 1268  
 Qy 6291 CCCCGTACTATT-----AGACCAGGCTGAGTAAACCAAGCTATCCT 6330  
 Db 1269 CCGCGGACTCATTAAGCCGCCGCCAGAAACCAAGTACTGGTGGCAATGACACGCTCCCT 1328  
 Qy 6331 TTAATAAACCATACGTGGAGTTTGTAAAGTAAACAGCT-CGTGTATATTTAGACGCCCTG 6389  
 Db 1329 TTAATAAAGTCAACCTTACTCCGCAAGGGGTAGTCTGTGTGGAATPACTGTCCAGCAGC 1388  
 Qy 6390 TCATAAAATCGATACCGACACCAATCAGGTAAATTAACCTTTTTTATTCAGAACTCAGGA 6449  
 Db 1389 CACAAATAATGGCCAGATGACAAAGTAAAGATCGACC--TTTATTGTATAGTAAACA 1446  
 Qy 6450 GGGCGTGTATACAGTGTGTAGTGGAGCATATCTGCCAATGGGAATGGCTGTATGAAT 6509  
 Db 1447 ATGCGTGTATACAAATGGTGTAGTGGAGCAGAGTTTCGCCAAGCTCTACGTCCGAACAGT 1506  
 Qy 6510 CCACACTAGTGTCCGATCAGGTGTCTGTGTACTTTTACTAGTCCGATGAACCTTAAT 6569  
 Db 1507 CGGTGTACAGGCTCTTATTAAGTGTTCGGTGTACTTGACCAAGCCGCGAACCTAGGT 1566  
 Qy 6570 TTGCGGTGTAGTTCGTACCAACCAAAACCTGCCCTGCCCTTTTAGAGCTCTA 6629  
 Db 1567 TGGTCTGTACAGTCTGTACAGCAAAAGGATCGGCGGTGCTTTTTCAGGAGTTA 1626  
 Qy 6630 TGCAGCTTTAATCGTGTCAAATAGCTTGTCTTTAAATGCAACTCTGTATCATCTTTA 6689  
 Db 1627 GGGACGTGTGATATATGACAGCTTCTGTCTGTAAATGACCGTGTATCATCTGAA 1686  
 Qy 6690 TGACAGTGTTCCTCAAAAAACAAAGATTTAAACACAGTTAAACTCGGTCTCGTGTAGT 6749  
 Db 1687 CGACAGCTGTCCAAAAACAAAGTTTCAGCTGCACGTTAAATCTGTATCTGAAAGT 1746  
 Qy 6750 CTTCTGTATTAACCTCTCAACAGAAATAATTTTAAACCAACTCGCAGTACTGAA 6809  
 Db 1747 CTTCTGTAAATGACAGTTTCTACCAAGAAACCTTTTACACGCTGCCCATCCACTGAA 1806  
 Qy 6810 ACAGTGTACGACAGTGTGGAAGTGTCTATCATCTCCTTCACTCCCTTTTGAAGG 6869  
 Db 1807 AGGAGGAGACAGCTCCCGTTGTGCGTTGTAGGATATCCCTAACTTCGGAGCGGAGAC 1866  
 Qy 6870 TTTTGGTTACACGACCATGTGAGATAGGACACAGCTCAGGGGGAGTCCGATCTGTAG 6929  
 Db 1867 GGGCGGAGCTCCCAAAAATGGGAGGACCACTCTGTGCACTCCCGGTCTGGGGTT 1926  
 Qy 6930 GGGGTAAATTCGTGGGCGTGGGGTACTCCAGTTTATAGATCTCCGCGAGACGCTA 6989  
 Db 1927 CTGATTCAGGGGCGCGGTGGGGGTTATGGAGAGTCAAAACTCTGGGAGTCCCTTAA 1986  
 Qy 6990 CGAGTTCTTCGTCAAACTCTACGAGCCCGCATCCACAGGGGAGTACTGAGTTAATAA 7049  
 Db 1987 TGAGCTCTCTCAAAACCTATGACGACGCTCCACTAGTGGCAGCATGCCGTTAATAA 2046  
 Qy 7050 GTCCCGAATTCATCTGTTCTGTTATTTTGGAGCTCTTTAGACATTAACCATGTTGC 7109  
 Db 2047 CACCCCTTATCTGTGCTGCTCAAGTTTGTACAACTGCTGCAAGGAATAGCCAAATTC 2106  
 Qy 7110 TCCGACACTGGCGTACGATGAACCGCTCTCTCTGTGTGTGATGACACCTGGAGTA 7169  
 Db 2107 CCTAGCCGCGGAACAGGTACGGCTCGCTTTGTGCTGTGACCAATATCTGAATGG 2166  
 Qy 7170 TCTTAGACAGTATAGGCCCGTTCCTACTTAAGGGTTGCCGCTTGTGTAGAGCGGTTGC 7229  
 Db 2167 TCTTTGCAAGGTATAGGCTCTTCTCAAGTTTATAGCGGGTACGTGGAGTCTGGATTGA 2226  
 Qy 7230 GGGCATTCAGCTCAGTGTGTCTATTTCCAGCATATCATCCGAGTGGCGGATAGTGA 7289  
 Db 2227 GGGTGGGACGAGGAGGTATCTAACTCTGAAGTATCTGATCCAGGACGGGTAAATGAT 2286  
 Qy 7290 ACTTGTACATATGATTTTAAAGATGGTCTTTTCCCGTAATGGCGACTCATGGTCTGTGTTT 7349

Db 2287 ACCTAAACAGATGTTTGAACAGGTGATCTTTTAAAGGGCCCTTCTCGATGTCTATGAAAA 2346  
 Qy 7350 A-----AGTTTACAGGGTGTAAATAAACTCAGAGGTCTCTCGCTGGACACTTT 7398  
 Db 2347 CTATGACACGCACTCTCTCTTAGGGTAAAGAGTTTCGGCGTCTCTGTGTGGAAGCTT 2406  
 Qy 7399 TTGCGGACTCTCCGAAACCGCTCAGGGATAAGCTACCAACGCGGTGCTTCTTTATAAACGA 7458  
 Db 2407 CGTGGGCTCTCGAGCGAACTGAAGGCCAACTCTACAGTGTGTCTCTTATAAATGA 2466  
 Qy 7459 AGCATAGTCAAAAAAATGACCCACGCTGACCTTAATAAGTCTACAGAGAGTCTCTCG 7518  
 Db 2467 CGCATACGAAACAATCTACGATCCAGTACCTTAATAGAGTGGTGAAGATGTGTGAT 2526  
 Qy 7519 ATGTATTGTATGAAATGTGCTCCCTGGGTCCCATCTCCGGGCTCATCGCGGATCTGAA 7578  
 Db 2527 TCGGATATGAAGAAATGTTTCAAGCTTGTGCGCTATGTGTCTGTCTTACAGACATTAA 2586  
 Qy 7579 TCTCTTAACTCTGTGTCTCTACCGGGTTCGGGGTTAAACCCGCGGGCGGCGAC 7638  
 Db 2587 CATGTTAACTTCTGTCTTTTTCGTCTCGAATGAGGACCAAGAGCGCGGCGG 2646  
 Qy 7639 CTGTAAACGTCGCGCGCAGAAATCGCCAGGGCATCTGTGAGAAATCTCACCGAGAGGGC 7698  
 Db 2647 GTAACAAGTGCATGCGCAGAGGCATCCCAAGGCATTTATTCGGATCTCTACGGAGAGAT 2706  
 Qy 7699 CTTTGTGTACGAAAGATGTTTATAGCTGTCTCTGCTGAGCGGGTCTGTATACCGCC 7758  
 Db 2707 CTTATTTCTGACAGAAAGGATTTCTGACAGCGCATGACAGCGGGTGAACCTGCTCTCC 2766  
 Qy 7759 CCAAGTGGCCAGAGTTCTGACAGCATCTACCGGAAATGAAGCCAAAGTGTTTGGGGGC 7818  
 Db 2767 AGCCATATGTAAGTACTACAGAAATATACACTGAAATGAAGCCAAATGCTTGGGGGC 2826  
 Qy 7819 GTGGCGTCTCTCATATGTTGACGGCGCTTATCATGGCTATGTCGACTCGTCTCTGT 7878  
 Db 2827 CTGAGGCGACTCTGCTGCAATCGAGGCCATATGATATTAACCTCTCTCTACTGAA 2886  
 Qy 7879 CACTTATTAATACCTGAGTGCAGGGGAAACTGGAACCTTAGACTCAAGACACTGTGAA 7938  
 Db 2887 GCTCTACAACAGTACGATACCGCGGCTCTCTCTGAGCAGTCCAGGGCCCTCTGCT 2946  
 Qy 7939 ACTTCTTTTCAACCCATCTTTTCTTCAAGAAATCTTAGCCCTATGACGCTACTGGCAA 7998  
 Db 2947 TTTGTTTTCDAACCGGCTTACCTTCCGAGGATATGSCGCGCTGGAGATCATGACCAA 3006  
 Qy 7999 CGGGAAGATGTTTCTGACAAATTTTACCATCAACGCTAGCGCGGAGAGAGCGGCC 8058  
 Db 3007 GGGTCAGCTCGCCCTGAAACTTTTACAGCATCACCGGTTCTGCTGAGAAACGCGCGCC 3066  
 Qy 8059 TGTCTGACTGGAAGTACTAGCGGATGACGTGTCGGGAAAGCAGCTTGTCCCGGATTC 8118  
 Db 3067 AATTACCGCGCAAGGTCTAGGACTGTCTTATCCAGGAAGCGGTCTCATGCCAGAAATC 3126  
 Qy 8119 CTTAATCTCTCCAGTATCGAAACCGGGCTTCTCCCGGCAACCCCTGGTTGACCTAGTAA 8178  
 Db 3127 TTTAATTTTGCCTTCTGAGCCAGGACTGTTGCGGCTTCCATGTTAGAGCTCAGCGA 3186  
 Qy 8179 TGTCTTAGAAATCCAGAAATCATCTCAGCGGCCCAACCCCTGAGTCAATTTGTCTATCAC 8238  
 Db 3187 TGTGTGGGAAACCGCGGTTTCTGAGCGCCCTCTGCTGAGCCAGTCTGTCTATTAG 3246  
 Qy 8239 AAACACGACCCAGTCTGCTCTAGTCAAGTCAAGTATTTAGCCCAACCCAGGGGCTGT 8298  
 Db 3247 CAAACCCATCCCAACATGCCGACACCGTCAAGCATCATCCCTTTTAAACC---CATCGG 3303  
 Qy 8299 TCCGGCCATGTTTATGACACGTGGAAACGGGTGTACAGAGCATTCACACAGAGC 8358  
 Db 3304 TACAGACCGCGGTTTATTAGTACGTGGCGGCGGCTCACAGATATGTTGTATCAACAC 3363  
 Qy 8359 ACAGAC 8364

Db 3364 ATCCAC 3369

RESULT 6  
AAV62153/c  
ID AAV62153 standard; DNA; 7361 BP.  
XX AC AAV62153;  
XX DT 18-DEC-1998 (first entry)  
XX DE  
XX XX HSV-2 strain SB5 Contig ID 1 DNA sequence.  
XX KW HSV-2 strain SB5; immunological response induction; therapy;  
KW antiviral identification; viral protein inhibitor; ss.  
XX OS Herpes simplex virus 2.  
XX PH Location/Qualifiers  
XX Key 322..864  
XX CDS /tag= a "ORF#1 protein"  
FT /product= "ORF#1 protein"  
FT /notes= "encoded protein shown in AAW72089"  
CDS 737..1741  
FT /tag= b  
FT /product= "ORF#2 protein"  
FT /notes= "encoded protein shown in AAW72090"  
CDS 1796..2497  
FT /\*tag= c  
FT /product= "ORF#3 protein"  
FT /notes= "encoded protein shown in AAW72091"  
CDS complement(2697..3302)  
FT /\*tag= d  
FT /product= "ORF#4 protein"  
FT /notes= "encoded protein shown in AAW72092"  
CDS complement(3369..6014)  
FT /\*tag= e  
FT /product= "ORF#5 protein"  
FT /notes= "encoded protein shown in AAW72093"  
CDS 5584..7361  
FT /\*tag= f  
FT /transl\_except= (pos: 7360..7361, aa: Arg-Val)  
FT /product= "ORF#6 protein"  
FT /notes= "encoded protein shown in AAW72094"  
XX WO9820016-A1.  
XX PD 14-MAY-1998.  
XX PF 31-OCT-1997; 97WO-US020016.  
XX PR 04-NOV-1996; 96US-0030279P.  
XX PR 09-JUN-1997; 97US-0049018P.  
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.  
XX ES Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB,  
PI Leary JJ;  
XX WPI: 1998-286847/25.  
XX DR P-PSDB; AAW72089, AAW72090, AAW72091, AAW72092, AAW72093, AAW72094.  
XX PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and  
PT treatment of infection or inducing immunological response in mammal.  
XX Claim 1; Page 338-341; 748pp; English.  
XX CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA  
CC sequence of the invention. This sequence was isolated from HSV-2 strain  
CC SB5 (deposited as ATCC VR-2346), is designated Contig ID 1, and encodes 6  
CC HSV-2 proteins. The proteins can be used for the treatment or prevention  
CC of disease, to induce an immunological response in a mammal or to  
CC identify inhibitors, activators or novel antivirals. Antagonists of the



```
RESULT 7
AAV62176/c
ID  AAV62176 standard; DNA; 117213 BP.
XX  AC  AAV62176;
XX  DT  13-JAN-1999 (first entry)
XX  DE  HSV-2 strain SB5 Contig ID 15 DNA sequence.
XX  KW  HSV-2 strain SB5; immunological response induction; therapy;
XX  KW  antiviral identification; viral protein inhibitor; ss.
XX  OS  Herpes simplex virus 2.
XX  FH  Location/Qualifiers
XX  CDS  755..1297
XX  FT  /tag= a
XX  FT  /product= "ORF#1 protein"
XX  FT  /note= "encoded protein shown in AAW72170"
XX  CDS  1170..2174
XX  FT  /tag= b
XX  FT  /product= "ORF#2 protein"
XX  FT  /note= "encoded protein shown in AAW72171"
XX  CDS  2229..2930
XX  FT  /tag= c
XX  FT  /product= "ORF#3 protein"
XX  FT  /note= "encoded protein shown in AAW72172"
XX  CDS  complement(3130..3735)
XX  FT  /tag= d
XX  FT  /product= "ORF#4 protein"
XX  FT  /note= "encoded protein shown in AAW72173"
XX  CDS  complement(3802..6447)
XX  FT  /tag= e
XX  FT  /product= "ORF#5 protein"
XX  FT  /note= "encoded protein shown in AAW72174"
XX  CDS  6017..8482
XX  FT  /tag= f
XX  FT  /product= "ORF#6g protein"
XX  FT  /note= "encoded protein shown in AAW72250"
XX  CDS  6026..8482
XX  FT  /tag= g
XX  FT  /product= "ORF#6f protein"
XX  FT  /note= "encoded protein shown in AAW72249"
XX  CDS  6085..8482
XX  FT  /tag= h
XX  FT  /product= "ORF#6e protein"
XX  FT  /note= "encoded protein shown in AAW72248"
XX  CDS  6167..8482
XX  FT  /tag= i
XX  FT  /product= "ORF#6d protein"
XX  FT  /note= "encoded protein shown in AAW72247"
XX  CDS  6296..8482
XX  FT  /tag= j
XX  FT  /product= "ORF#6c protein"
XX  FT  /note= "encoded protein shown in AAW72246"
XX  CDS  6326..8482
XX  FT  /tag= k
XX  FT  /product= "ORF#6b protein"
XX  FT  /note= "encoded protein shown in AAW72245"
XX  CDS  6446..8482
XX  FT  /tag= l
XX  FT  /product= "ORF#6a protein"
XX  FT  /transl_except= (pos: 7400..7402, aa: Ala-Ala)
XX  FT  /transl_except= (pos: 7481..7486, aa: Ile)
XX  FT  /note= "encoded protein shown in AAW72244"
XX  CDS  8457..9347
XX  FT  /tag= m
XX  FT  /product= "ORF#7 protein"
XX  FT  /note= "encoded protein shown in AAW72175"
XX  CDS  complement(9604..11855)
XX  FT  /tag= n
XX  FT  /transl_except= (pos: 11635..11636, aa: Ala)
XX  FT  /product= "ORF#8 protein"
XX  FT  /note= "encoded protein shown in AAW72176"
XX  CDS  complement(11905..14520)
XX  FT  /tag= p
XX  FT  /product= "ORF#9a protein"
XX  FT  /note= "encoded protein shown in AAW72223"
XX  CDS  complement(11905..14508)
XX  FT  /tag= o
XX  FT  /product= "ORF#9b protein"
XX  FT  /note= "encoded protein shown in AAW72222"
XX  CDS  14399..15802
XX  FT  /tag= q
XX  FT  /product= "ORF#10 protein"
XX  FT  /note= "encoded protein shown in AAW72177"
XX  CDS  complement(15996..16286)
XX  FT  /tag= x
XX  FT  /product= "ORF#11 protein"
XX  FT  /note= "encoded protein shown in AAW72178"
XX  CDS  complement(16202..18064)
XX  FT  /tag= s
XX  FT  /product= "ORF#12 protein"
XX  FT  /note= "encoded protein shown in AAW72179"
XX  CDS  complement(18105..19661)
XX  FT  /tag= t
XX  FT  /product= "ORF#13 protein"
XX  FT  /note= "encoded protein shown in AAW72180"
XX  CDS  complement(19415..20074)
XX  FT  /tag= u
XX  FT  /product= "ORF#14 protein"
XX  FT  /note= "encoded protein shown in AAW72181"
XX  CDS  20155..21453
XX  FT  /tag= v
XX  FT  /product= "ORF#15 protein"
XX  FT  /note= "encoded protein shown in AAW72182"
XX  CDS  complement(21326..22291)
XX  FT  /tag= w
XX  FT  /product= "ORF#16 protein"
XX  FT  /note= "encoded protein shown in AAW72183"
XX  CDS  complement(22546..24654)
XX  FT  /tag= x
XX  FT  /product= "ORF#17 protein"
XX  FT  /note= "encoded protein shown in AAW72184"
XX  CDS  24684..25955
XX  FT  /tag= y
XX  FT  /product= "ORF#18 protein"
XX  FT  /note= "encoded protein shown in AAW72185"
XX  CDS  complement(26295..27251)
XX  FT  /tag= z
XX  FT  /product= "ORF#19 protein"
XX  FT  /note= "encoded protein shown in AAW72186"
XX  CDS  complement(27630..31784)
XX  FT  /tag= ab
XX  FT  /product= "ORF#20a protein"
XX  FT  /note= "encoded protein shown in AAW72224"
XX  CDS  complement(27630..31754)
XX  FT  /tag= aa
XX  FT  /product= "ORF#20b protein"
XX  FT  /note= "encoded protein shown in AAW72225"
XX  CDS  complement(32067..32735)
XX  FT  /tag= ac
XX  FT  /product= "ORF#21 protein"
XX  FT  /note= "encoded protein shown in AAW72187"
XX  CDS  33140..34984
XX  FT  /tag= ad
XX  FT  /product= "ORF#22a protein"
XX  FT  /note= "encoded protein shown in AAW72226"
XX  CDS  33386..34984
XX  FT  /tag= ae
XX  FT  /product= "ORF#22b protein"
XX  FT  /note= "encoded protein shown in AAW72227"
XX  CDS  complement(35205..37721)
```



PN WO200176643-A1.  
 XX 18-OCT-2001.  
 PD 06-APR-2001; 2001WO-US011372.  
 XX 07-APR-2000; 2000US-0195680P.  
 PF (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX Orson FM, Kinsey BM, Bhogal BS;  
 XX WPI; 2002-066308/09.  
 XX  
 XX Composition for oral delivery of vaccines, comprises expression vector  
 XX containing antigenic genomic sequence, bound to aggregated protein-  
 XX polycationic polymer conjugate or suspension.  
 XX  
 XX Disclosure; Page 90-132; 145pp; English.  
 XX  
 XX The invention relates to a composition comprising an expression vector  
 XX bound to an aggregated protein-polycationic polymer conjugate or  
 XX suspension. The expression vector contains a promoter polynucleotide  
 XX sequence operatively linked to a polynucleotide sequence encoding an  
 XX antigen which is a fragment of a gene or genome associated with an  
 XX infectious disease, cancer and autoimmune disease such as rheumatoid  
 XX arthritis, vasculitis, and multiple sclerosis, pathogenic genomes  
 XX consisting of bacterium, fungus, protozoa and virus such as human  
 XX immunodeficiency virus (HIV) herpes simplex virus (HSV), hepatitis C  
 XX virus (HCV), influenza and respiratory syncytial virus (RSV), and  
 XX optionally comprising a nucleotide sequence encoding a cytokine (or a  
 XX cytokine expression vector), is useful for inducing an immune response  
 XX (systemic and/or mucosal) in an organism. The cytokine expression vector  
 XX contains a sequence for granulocyte macrophage-colony stimulating factor  
 XX (GM-CSF) and interleukin-12 (IL-12). The polynucleotide sequences encoding  
 XX the antigen and the cytokine are under transcriptional control of same or  
 XX different promoter polynucleotide sequences. The expression vector, as a  
 XX DNA vaccine is useful for treating a condition in an organism. The  
 XX present sequence is human herpesvirus 2 complete DNA genome related to  
 XX the invention  
 XX  
 XX Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 U; 0 Other;  
 XX  
 XX Query Match 2.2%; Score 200.8; DB 6; Length 154746;  
 XX Best Local Similarity 52.3%; Pred. No. le-46;  
 XX Matches 495; Conservative 0; Mismatches 442; Indels 9; Gaps 2;  
 XX  
 XX 585 CAATATTATTGTAATCGACGAGGCGGGAACCTCTCTTACATACCTACCCCGTCGT 644  
 XX 14524 CAACATATCTGTCATCGACGAGGCGGGCTCTCGGGCGGCACCTACTCAGACCGTGGT 14465  
 XX 645 GTTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGGCGGT 704  
 XX 14464 GTATTGCTGGTGGATGATTACGCCCTGTACCAACACCCCGCCAGTACCGGGCGCGCTGGC 14405  
 XX 705 TCCGTGCATAGTATGTCGTCGTCGCCAAGCAGACAAACCGCTTCCAGTCAACTTACAA 764  
 XX 14404 GCCGTGCTGGTGGTGGGCTGCGGACCAACGCGCTGCTGGAGTCCACCTTCA 14345  
 XX 765 CCACGGGACGCAAAAGACGAGATATGTCGTGGAGAAATCTTAATCATGATGAGG 824  
 XX 14344 ACACCAAAATGTCGATCTCGTCCGCGCAGAGCGAAACGCTCAGTACCTCATCTG 14285  
 XX 825 GAAGAAGTCTGATCTGATGATGATACCTTGGAGAGAACTGGGCGCTGTTTATAACAA 884  
 XX 14284 CAACCGACCTTACCGGAGTACCGGCTCTCGACAGCTGGGCCATTTTCAATACAA 14225  
 XX 885 CAAGCGCTGCACCGATTCGAGTTTCGGACACTTCTCTAAATTTTAGATATATCTTCC 944  
 XX 14224 CAAGCGATGTGGGAGCAGGAGTTCGGAACTCATGAAGGTGCTGGAGTACGGCCCTCC 14165  
 XX 945 CATTCCTGACGAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004

Db 14164 CATCACCGAGGAGCACATGTCAGCTTTGTGGACCGCTTTTGTCTCCGGAAGTTTACATCAC 14105  
 QY 1005 GGATCCCTTTAGAAATACATTTGGCTGACCCGCTCTTTTGTGTCATAGCGAGGTAAAGGC 1064  
 Db 14104 CAACCGGCGCAACCTTCGGGGTGGACGGGCTGTTCTGCTCCACAGAGGTCAGCGC 14045  
 QY 1065 GTATCTAAACAACTACACATGCTAAACGCTAGGGGGCGATACACAGGGACAGAAAGCT 1124  
 Db 14044 GTACATGCGCAAGCTCCAGCGCTACCTAAAGGTGACTCGCGA---GGGGGAGTTTGTGT 13988  
 QY 1125 CTTTACCTGTCCTGGTGTGTCGAGGCTTTGTGTAAGCCGTTTTCAGGAATACAAAGCGGC 1184  
 Db 13987 GTTTACCTCTCCCGTGTCTACGTTTGTGTCGTTCAAGAGTTTTCAGGATATCAGCGCT 13928  
 QY 1185 CGTCAACCTCACCAACCTCACCGTACCGTACCGTGGTGGTGCACAAAACCTCTTTAAAGTTAAG 1244  
 Db 13927 CAGCGAGCAACCCAGCTGACCATGGAAAAGTGGATCAGCGCCAAACGCACTCGCATCAC 13868  
 QY 1245 TAATTTATTCGAGTTTGTGGACCAAGGACATGTCATAGTCCCGGCAATC-----CAC 1298  
 Db 13867 CAATCTACTCCAGAGTCAGGACCAAGGACCGCGGCGACGTCGCTGTGAGGTGCACAGCA 13808  
 QY 1299 TGAACGTTTCGACCCAGGTTTACCTTTATCACCAAGTTTGTAAACACAGCCAGCTGCTTTT 1358  
 Db 13807 GCAACAGTAGTGTGTCGGCCGGAACGACATCACGTACGTCCTCAACAGCCAGGTGCGGT 13748  
 QY 1359 AAATCGAAAAACAAAAAATGCATATGTCGGTTTCAGGTTACGTTTGTGAGTTCAAAAG 1418  
 Db 13747 GACCGCGCGCTCCGAAAGATGGTTTGGGTTTCGACGGGACGTTTCGACCTTCGAGGC 13698  
 QY 1419 AATCTAGACAGGCACTCTTCGTGGACAGCATTCGCAAGACCGCTCCGGAATAGTATA 1478  
 Db 13687 TGTGTCGGGACGACACCTTCGTGAACACCCAGGGGAGACCTCGGTGGAGTTGCGCTA 13628  
 QY 1479 CGGTTTCTTAAACACGCTACTGTACAAACGCAATGATCTCGTTTTCAC 1524  
 Db 13627 CCGGTTCTGTCGCGGCTCATGTTCCGGCGGCTGATTCACTTTTAC 13582  
 XX  
 XX RESULT 9  
 XX AAV62130  
 XX ID AAV62130 standard; DNA; 8952 BP.  
 XX AC AAV62130;  
 XX AC AAV62130;  
 XX DT 07-DEC-1998 (first entry)  
 XX DE HSV-2 strain SB5 Contig ID 100 DNA sequence.  
 XX XX  
 XX HSV-2 strain SB5; immunological response induction; therapy;  
 XX KW antiviral identification; viral protein inhibitor; ss.  
 XX OS Herpes simplex virus 2.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS complement (1..1802)  
 XX /\*tag= a  
 XX /product= "ORF#1 protein"  
 XX /transl\_except= (pos: 1..2, aa: Ala-Xaa)  
 XX /note= "Xaa= unspecified amino acid"  
 XX 1372..4017  
 XX CDS /tag= b  
 XX /product= "ORF#2 protein"  
 XX 4084..4689  
 XX /tag= c  
 XX CDS /product= "ORF#3 protein"  
 XX complement (4889..5407)  
 XX /tag= d  
 XX CDS /product= "ORF#4 protein"  
 XX complement (5646..6650)  
 XX /tag= e  
 XX CDS /product= "ORF#5 protein"  
 XX complement (6523..7074)



CC control of vaccines for use in the prevention of poxvirus diseases such  
 CC as chickenpox  
 XX  
 SQ Sequence 125157 BP; 33776 A; 28439 C; 28285 G; 33624 T; 0 U; 33 Other;  
 Query Match 1.7%; Score 154; DB 5; Length 125157;  
 Best Local Similarity 56.2%; Pred. No. 9e-33;  
 Matches 289; Conservative 0; Mismatches 225; Indels 0; Gaps 0;  
 QY 585 CAATATTATTGTAATCGACGCGCGGACCGCTGCTCTTACATATCTACACCGCTCGT 644  
 Db 96652 CAACGTTATTGTTATGATGAAGCAGGATGCTAGGCGGTATATTCTCAGCGCGTGT 96711  
 QY 645 GTTCTTTTACTGCTTCTTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGCGGCGGT 704  
 Db 96712 TTACTGTTGCTGGCTTTTGAATGCTATATATCAAAAGCCCTCAGTACATAAAACGTCGAAA 96771  
 QY 705 TCCGTGCATAGTATGCTGCTGCTTCCCAACGACGACGACGCTTCCAGTCACTTACAA 764  
 Db 96772 ACGGTCATAGTATGCTGCTGCTTCCCAACGACGACGCTTCCAGTCACTTACAA 96831  
 QY 765 CCACGGGACGCAAAAGACGAGATATCTGCTGGAGAACATCTTAACATTCATGATGAG 824  
 Db 96832 ACATGACATGACGCTTCCACAGCTAACTCCTAGTGAATAATATCTCAGTATATATCTG 96891  
 QY 825 GAAGAAGTCTGATCTGAGTACGTACCTTGGAGAGAACTGGCGGCTGTTTATTAACAA 884  
 Db 96892 CAATCAACTCTGCTCAATATATCAATCTCACTACCTGGCAATCTTATTAATAA 96951  
 QY 885 CAAGCGCTGACCGATCTGAGTTCGGACACCTTCTAAATTTTACGATATATCTTC 944  
 Db 96952 CAACGATGTCAGAGGACGATTTTGGAAATCTTTTAAACGCTTGGTACGGGTACC 97011  
 QY 945 CATTCTGACGAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004  
 Db 97012 TATTACGAGACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 97071  
 QY 1005 GGATCCTTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1064  
 Db 97072 CAATCCTGCTAATCTCCCGGATGACGCGTCTGATTCGTTCGATTAAGAGGTGAGCG 97131  
 QY 1065 GTATCTTAAACACCTTACACATGCTTAACGCTA 1098  
 Db 97132 GTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97165  
 RESULT 11  
 ID AAT85076  
 XX AAT85076 standard; DNA; 2871 BP.  
 AC AAT85076;  
 XX  
 DT 18-MAR-1998 (first entry)  
 DE Human cytomegalovirus (HCMV) UL105 open reading frame.  
 XX  
 KW Human cytomegalovirus primase; HCMV UL105; open reading frame; orf;  
 screening; inhibitor; infection; diagnosis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Location/Qualifiers  
 FT 1..2871  
 FT /\*tag= a  
 FT /product= "HCMV\_UL105\_orf\_translation\_product"  
 XX  
 FN GB2311069-A.  
 XX  
 PD 17-SEP-1997.  
 XX  
 PF 05-MAR-1997; 97GB-00004577.  
 XX  
 PR 14-MAR-1996; 96US-0013546P.

PR 04-APR-1996; 96GB-00007117.  
 XX (MERI ) MERCK & CO INC.  
 XX  
 PI Gotlib L, Hazuda DJ, Lafemina RL;  
 XX WPI; 1997-427907/40.  
 DR P-PSDB; AAW27086.  
 XX  
 XX Novel human cytomegalovirus primase - used for screening for HCMV  
 PT antivirals and in diagnosis of HCMV related diseases.  
 XX  
 PS Example 1; Page 20-22; 28pp; English.  
 XX  
 CC The present sequence represents the human cytomegalovirus (HCMV) UL105  
 CC open reading frame (orf) spanning nucleotides 151926 through 154793 of  
 CC HCMV AD169 which was amplified by PCR primers AAT85074-5. The amplified  
 CC product was digested with BglII and EcoRI and cloned into BglII/EcoRI  
 CC digested p817 vector DNA to yield the plasmid p88717 UL105. The  
 CC amplified HCMV UL105 orf sequence and its translational product were  
 CC determined. The HCMV UL105 orf was recombined into the baculovirus AcNPV  
 CC genome by standard co-transfection protocols. The invention relates to a  
 CC new human cytomegalovirus (HCMV) primase. A screening assay for compounds  
 CC which inhibit HCMV primase (preferably with an IC50 of not greater than  
 CC 200nM), comprises incubation of the compound with the primase. The  
 CC primase of the invention is useful as a screening tool for HCMV  
 CC antivirals, and as a diagnostic tool for diseases resulting from HCMV  
 CC infection  
 XX  
 SQ Sequence 2871 BP; 543 A; 971 C; 795 G; 562 T; 0 U; 0 Other;  
 Query Match 1.7%; Score 153.4; DB 2; Length 2871;  
 Best Local Similarity 50.3%; Pred. No. 8.5e-34;  
 Matches 376; Conservative 0; Mismatches 371; Indels 0; Gaps 0;  
 QY 585 CAATATTATTGTAATCGACGCGCGGAAACCCCTGCTCTTACATCTCACCACCGTCTG 644  
 Db 765 CAATATCATCTCATCGACGAGTGGCCCTTATGCTGCGCTACATGCTGAGGTGGTGT 824  
 QY 645 GTTCTTTTACTGCTTCTTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGCGGCGGT 704  
 Db 825 GTTTTCTTACTCTTTTACACGCGCTGGCGACACGCGACTTTACGCGAACCGCGCT 884  
 QY 705 TCCGTGCATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764  
 Db 885 GCCCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944  
 QY 765 CCACGGGACGCAAAAGACGAGATATCTGCTGCGAGAACATCTTAACATTCATGATAG 824  
 Db 945 CCATCTACGCAAAACAGAGCTGCGCAGGCGGTGACGCTCTCTCGGCGCTGATTCA 1004  
 QY 825 GAAGAAGTCTGATCTGAGTACGTACATCTTGGAGAGAACTGGCGGCTGTTTATAAACAA 884  
 Db 1005 GAACGAGTGTCTCATCACTACTGCGACATCGCCGCAACTGGGTCTATGTTTATTCACAA 1064  
 QY 885 CAACGCTGACCGATCTGCGAGTTCGACACCTTCTAAATAATTTTAGAGTATAATCTTC 944  
 Db 1065 CAACGCTTGCACCGACCTGCTGCGACCTTTTGGGACCTGCTCAAGTACATGAGTTCGGTATCC 1124  
 QY 945 CATTCTGACGAGTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004  
 Db 1125 GCTCAGGAGGAGCAGCTGGCCCTACGTGATCGCTTCTGCGGCGCCCGAGCTCCATCCG 1184  
 QY 1005 GGATCCTTTTAAATATCAATTTGGCTGGACCGGACTCTTTTGTACATAGCGAGTAAAGGC 1064  
 Db 1185 CAACCCCTCTGACCGCCCGAGATGACGCGGCTTTTCTCTCACAGTCTGAGGTGAGGC 1244  
 QY 1065 GTATCTTAAACAACTTACACATGCTTAAGCTAGGGGGGATACACGAGGACACGAGCT 1124  
 Db 1245 TTACTTCAAGCGGCTGCACGAGCAGATCCGCTGACGCGCCACCGCTCTCTTTGATCT 1304  
 QY 1125 CTTTACCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1184

Db 1305 GCCGCTCTACTGGTCAACAACCGCGGTACCGAGGCTCTGCGAGCTGGCCGACCC 1364  
 QY 1185 CGTCAACCTCACCACCTCAGCTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAG 1244  
 Db 1365 GCTGGGCGACTCGCCGAGCCGCTCGAGCTCTGCTCCGCGAGAACTTGGCGGATCAT 1424  
 QY 1245 TAATTATTGCGCACTTTGTGACGAGGACATGTCCATAGTTCGCCACGCAATCCACTGAACG 1304  
 Db 1425 TAATACTCGCACTTTGTGACGAGGACATGTCCATAGTTCGCCACGCAATCCACTGAACG 1484  
 QY 1305 TTGACCCGAGTTACCTTTATACCAA 1331  
 Db 1485 CCCCGCGCGCGAGCTGCTGGTCCACCAA 1511

RESULT 12

AAT85062

ID AAT85062 standard; DNA; 2871 BP.

XX AC AAT85062;

XX 18-MAR-1998 (first entry)

XX Human cytomegalovirus (HCMV) helicase.

XX Human cytomegalovirus helicase; HCMV UL105; open reading frame; orf;

XX screening; inhibitor; infection; diagnosis; antiviral; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..2871

FT FT /\*tag= a

XX /product= "HCMV\_helicase"

XX GB2311068-A.

XX 17-SEP-1997.

XX 05-MAR-1997; 97GB-00004575.

XX 14-MAR-1996; 96US-0013389P.

XX 04-APR-1996; 96GB-00007118.

XX (MERI ) MERCK & CO INC.

XX Gotlib L, Hazuda DJ, Lafemina RL;

XX WPI; 1997-427906/40.

XX P-PSDB; AAM27083.

XX Novel human cytomegalovirus helicase - used for screening for HCMV

XX antivirals and in diagnosis of HCMV related diseases.

XX Claim 2; Page 11-12; 28pp; English.

XX The present sequence represents DNA encoding human cytomegalovirus (HCMV)

XX helicase, spanning nucleotides 151926 through 154793 of HCMV AD169 which

XX was amplified by PCR primers AAT85083-4. The amplified product was

XX digested with BglII and EcoRI and cloned into BglII/EcoRI digested pB877

XX vector DNA to yield the plasmid pB877 UL105. The amplified HCMV UL105

XX orf sequence and its translational product, helicase, were determined. The

XX HCMV UL105 orf was recombined into the baculovirus AcNPV genome by

XX standard co-transfection protocols. This new HCMV helicase is used in a

XX screening assay for compounds which inhibit HCMV helicase (preferably

XX with an IC50 of not greater than 200nM), and as a diagnostic tool for

XX diseases resulting from HCMV infection

XX Sequence 2871 BP; 543 A; 971 C; 795 G; 562 T; 0 U; 0 Other;

XX Query March

XX Best Local Similarity 1.7%; Score 153.4; DB 2; Length 2871;

XX Matches 376; Conservative 0; Mismatches 371; Indels 0; Gaps 0;

QY 585 CAATATTATTGTAATGACGAGCGCGGAAACCTCTCTCTTTACATATCTACACCGCTGT 644  
 Db 765 CAATATTATCTGTCATGACGAGTGCAGGCTTATGCTGGCTACATGCTGAGGTGGT 824  
 QY 645 GTTCTTTTACTGTTCTTAACAGCTGCTTAATACCCCTCTTTACCGCCAGGCGGCT 704  
 Db 825 GTTCTTTTACTGTTCTTAACAGCTGCTTAATACCCCTCTTTACCGCCAGGCGGCT 884  
 QY 705 TCCGTGATAGTATGCTGCTGCTTCCCAACGACGACAAACGCGTTCCAGTCAATACAA 764  
 Db 885 GCTCTGATATCTGCTGCTGCTTCCCAACGACGACGAGCGCTGGAGAGCGCTAGCA 944  
 QY 765 CCACGGGACGCAAAAGAGCGAGATATGCTGCTGCGAGAAATCTCTTAATATGATAGG 824  
 Db 945 CCCTTACAGCAAAACGAGCGCTGCGCAAGGCGTTGACGCTCTCTGCGCGCTGATCA 1004  
 QY 825 GAAGNAGTCTGTAATCTGAGTACGTACACTTGGAGAGAAACCTGGCGCTGTTTATAACAA 884  
 Db 1005 GAACGAGGTCTCATCAACTACTGCGACATCGCGACATCGCGACATCGGTATGCC 1064  
 QY 885 CAAGCGCTGCAACCGATCTGCAAGTTCGGACACCTTCTAAATAATTTAGAGTATATCTTCC 944  
 Db 1065 CAAGGTTGCAACCGACCTGGACTTTGGCGACCTGCTCAAGTACATGAGTTCGGTATCCC 1124  
 QY 945 CATCTCTGACGAAGTCAATGAGTACGTAGACAGATTTGCTGTTCTTAAAGTAAAGTAT 1004  
 Db 1125 GCTCAAGGAGGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1184  
 QY 1005 GGATCTTTAGATATACATTTGGTCTGACCCGACTCTTTTGTACATAGCGAGGTAAGGC 1064  
 Db 1185 CAACCCCTCTGTAACCGCGCGAGATGACGCGCTTTTCTCTCACACGTCGAGGTGACGC 1244  
 QY 1065 GTATCTAAACAACTACACATGCTTAAACGCTAGGGGCGATACCGAGGACACGAGCT 1124  
 Db 1245 TTACTTCAAGCGGCTGCAACGAGCAGATCCGCTGAGCGAGCGCCACCGCTCTCTTGTATCT 1304  
 QY 1125 CTTTACCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1184  
 Db 1305 GCGCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1364  
 QY 1185 CGTCAACCTCAACCACTCACATGCTTAAACGCTAGGGGCGATACCGAGGACACGAGCT 1244  
 Db 1365 GCTGGCGACTCGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1424  
 QY 1245 TAATTATTGCGAGTTTGGGACCGAGCATGTCATAGTCGCCACGGAATCCACTGAACG 1304  
 Db 1425 TAATCTCTGCAAGTTTTCGACCACTCTCCAGGAGATCACCAGGAGCGCTGG 1484  
 QY 1305 TTGACCCGAGTTACCTTTATACCAA 1331  
 Db 1485 CCCCGCGCGCGAGCTGCTGGTCCACCAA 1511

RESULT 13

ACC70010

ID ACC70010 standard; DNA; 5290 BP.

XX AC ACC70010;

XX 29-JUL-2003 (first entry)

XX Nucleotide sequence of the CMV immediate early region transcript.

XX Cancer; hyperproliferative disorder; viral vector; H19 promoter;

XX carcinoma; sarcoma; adenoma; ganglioblastoma; bladder carcinoma;

XX immediate early region; CMV; ss.

XX Cytomegalovirus.

XX WO2003035883-A2.

XX 01-MAY-2003.



XX PF 22-OCT-2002; 2002WO-IL000843.  
 XX PR 22-OCT-2001; 2001US-00012131.  
 XX PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 XX PI Hochberg A, Ayesh S;  
 XX DR WPI; 2003-441263/41.  
 XX PT Treating cancer or hyperproliferative disorder comprises administering a  
 PT conditionally replicative viral vector, where the viral nucleic acid  
 PT encoding a product essential for its replication is operably linked to  
 PT H19 regulatory sequence.  
 XX PS Disclosure; Page 119-120; 122pp; English.  
 XX CC The specification describes a method of treating cancer or  
 CC hyperproliferative disorders. The method comprises administering a  
 CC conditionally replicative viral vector, where at least one of the viral  
 CC nucleic acids encoding a product essential for its replication is  
 CC operably linked to an H19 regulatory sequence. The method is useful for  
 CC treating cancer or hyperproliferative disorders. The cancer is selected  
 CC from many different types, for example, carcinoma, sarcoma, adenoma and  
 CC ganglioblastoma. Preferably, the cancer is bladder carcinoma. The present  
 CC sequence represents the CMV immediate early region transcript. It may be  
 CC used to produce vectors for use in the method of the invention  
 XX SQ Sequence 5290 BP; 990 A; 1753 C; 1534 G; 1013 T; 0 U; 0 Other;  
 Query Match 1.7%; Score 153.4; DB 8; Length 5290;  
 Best Local Similarity 50.3%; Pred. No. 1.3e-33;  
 Matches 376; Conservative 0; Mismatches 371; Indels 0; Gaps 0;  
 Qy 585 CAATATTATTGATTCGACGAGCGCGAACCCTCTCTTACATATCTACACCGTGT 644  
 Db 2690 CAATATCATCTATCGACGAGTGGCGCTTATCTCGGTACATGTCGAGGTGGT 2749  
 Qy 645 GTTCTTTTACTGGTCTTAACACTGGCTAAATACCCCTCTTTACCGCGAGCGGT 704  
 Db 2750 GTTCTTTTACTGCTTTTACACGCGCTGGGCGACACCGGACTTTTACCGGAAACCGCGGT 2809  
 Qy 705 TCCGTGATAGTATGCTGCTGCGCAACGACGACAAACGCGTTCAGTCAACTTACAA 764  
 Db 2810 GCCCTGATCATCTGCTGCTGCTGCGCCACGACGAGCGCTGGAGAGCGCTACGA 2869  
 Qy 765 CCACGGACGCAAAAGACGGAGATATGCTGCTGCGAGAAATCTTAACATTCATGATAG 824  
 Db 2870 CCACGTACGCAAAACAGACGCTGGCGCAACGCGGCTTACGCTCTCGCGCTGATTC 2929  
 Qy 825 GAAGAAGTCTATCTGAGTACGTACACTTGGAGAGAACTGGGCGCTGTTTATAACAA 884  
 Db 2930 GAACGAGTGTCTATCACTACTCGGACATCGCCGACACTGGGTCTATTTATTCACAA 2989  
 Qy 885 CAAGCGTGCACCGATCTGAGTTCGGACACTTCTAAATAATTTAGATATATCTTCC 944  
 Db 2990 CAAGCGTGCACCGACTTGGGACTTGGGACCTGCTCAAGTACATGAGTTCGATATCCC 3049  
 Qy 945 CATCTCTACGAGTCTAGTCTAGCTAGACAGATTTGCTTCTCTAAAGTAAAGTATAT 1004  
 Db 3050 GCTCAAGGAGGACGCTGGCTTACGCTGATCGTCTGCGGCGCCGCTCCATCCG 3109  
 Qy 1005 GGATCTTTTAGAATACATTTGCTGGACCCGACTCTTTTGTTCATAGCAGGTAAAGGC 1064  
 Db 3110 CAACCCCTCGTAGCGCGCGAGATGACGCGCTTTTCTCTCACATCGCAGGTGCAGGC 3169  
 Qy 1065 GTATCTACAACACTTACACATCTCTAACGCTAGGGGCGCATACACGAGGACGAGCT 1124  
 Db 3170 TTACTTCAAGCGGTGACGAGCAGATCCGCTTACGAGCGCGCCACCGTCTCTTTGATCT 3229  
 Qy 1125 CTTTACCTGTCCGCTGTGCGAGGTGTTTGTGAAGCGGTTTGAGGAATACAGCGGC 1184

Db 3230 GCCCGTCTACTGCTGTGTCAACACCGCGCTACCCAGGAGCTCTCGAGCTGGCGACCC 3289  
 Qy 1185 CGTCAACCTCACCACCTCAGCGTACCGAGTGGGTGACAAAACCTCTTTAAGTTAAG 1244  
 Db 3290 GCTGGCGACTCGCGGAGCGCGTCTGCTGTTCCGCCAGACTTGGCGCGCATCAT 3349  
 Qy 1245 TAAATTATTCGAGTTTGTGGACCGAGGACATGTCATATGTCGCCACGGAATCCTGTAACG 1304  
 Db 3350 TAACTACTCGCAGTTTGTGACCAACACCTCTCCAGCGAGATCACCAAGGAGCGCTCG 3409  
 Qy 1305 TTCGACCCAGGTTACCTTTATCAACAA 1331  
 Db 3410 CCCCAGCGCGACGCTGTTGCCACCA 3436  
 RESULT 14  
 ABQ74179  
 ID ABQ74179 standard; DNA; 229354 BP.  
 XX AC ABQ74179;  
 DT 13-OCT-2002 (first entry)  
 XX Human cytomegalovirus strain AD169 genomic sequence SEQ ID NO:64.  
 DE Human cytomegalovirus; HCMV; CMV; Yeast artificial chromosome; YAC;  
 KW vaccine; immunisation; infection; anti-inflammatory; immunosuppressive;  
 KW virulence; hepatotropic; congenital infection; jaundice; pneumonitis;  
 KW respiratory distress; convulsive seizure; mental retardation; hepatitis;  
 KW neurologic disability; asymptomatic adult infection; mononucleosis;  
 KW retinitis; acquired immune deficiency syndrome; AIDS; genome; ds.  
 XX Human cytomegalovirus.  
 OS  
 XX Key Location/Qualifiers  
 FH Promoter 80996..81441  
 FT /\*tag= a  
 FT /\*note= "UL54 promoter"  
 FT Promoter 128318..128750  
 FT /\*tag= b  
 FT /\*note= "UL86 promoter"  
 XX  
 FN WO200257437-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 30-OCT-2001; 2001WO-US047943.  
 XX  
 PR 03-NOV-2000; 2000US-00705400.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 XX  
 PI Ghazal P, Huang H;  
 XX  
 DR WPI; 2002-590735/63.  
 XX  
 CC Cytomegalovirus-yeast artificial chromosome (CMV-YAC) DNA molecule useful  
 CC as a vaccine for treating CMV infections e.g. jaundice, respiratory  
 CC distress, convulsive seizures, mononucleosis, hepatitis, pneumonitis and  
 CC retinitis.  
 XX  
 PS Claim 5; Page 64-131; 132pp; English.  
 XX  
 CC The present invention describes an isolated recombinant DNA molecule  
 CC comprising a yeast artificial chromosome (YAC) including at least a  
 CC portion of a human cytomegalovirus (HCMV) genome. Also described is a  
 CC composition suitable for use as a vaccine comprising the isolated  
 CC recombinant DNA molecule and an excipient. The YAC has anti-inflammatory,  
 CC immunosuppressive, virulence, and hepatotropic activities, and can be used  
 CC in vaccine production. The YAC vector and the recombinant DNA molecule  
 CC are useful as vaccines for treating infections caused by CMV, e.g.  
 CC congenital infections such as jaundice, respiratory distress and  
 CC convulsive seizures, which may result in mental retardation, neurologic

CC disability or death; or asymptomatic adult infections such as  
CC mononucleosis, hepatitis, pneumonitis and retinitis. The vaccine can also  
CC be used for immuno-compromised patients suffering from acquired immune  
CC deficiency syndrome (AIDS), or undergoing chemotherapy or tissue  
CC rejection therapy after organ transplantation. The vector is also useful  
CC as a tool for studying the replication of CMV using cell-line models and  
CC the function of essential cis acting genes or nucleic acid sequences. The  
CC YAC vector provides an ease and rapidity in introducing single or  
CC multiple mutations into a CMV DNA sequence. The capacity of the vectors  
CC to generate multiple mutations exceeds the capacity of other vectors,  
CC e.g. bacterial artificial chromosome. The present sequence represents  
CC human cytomegalovirus strain AD169 genomic sequence (also see Genbank  
CC X17403.1), given in the present invention  
XX  
SQ Sequence 229354 BP; 49475 A; 64911 C; 66192 G; 48776 T; 0 U; 0 Other;  
Query Match 1.7%; Score 153.4; DB 6; Length 229354;  
Best Local Similarity 50.3%; Pred. No. 2.1e-32;  
Matches 376; Conservative 0; Mismatches 371; Indels 0; Gaps 0;  
QY 585 CAATATATTGTAATCGACGAGCGCCGGAACCCCTGCTCTTACATACCTACACACGCGTGT 644  
Db 152690 CAATATCATGTCATCGACGAGTGGCGCTTATGCTGCGCTACATGCTGAGGTGTGT 152749  
QY 645 GTTCTTTTACTGTTCTCTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGCGGCGGT 704  
Db 152750 GTTTTTTACTACTTTTACACGCCCTGGCGACACGCGACTTTACCGGAAACGCGGT 152809  
QY 705 TCGGTGATAGTATGGTTCGCGCAACGCGACAAACGGTTCACGTCAACTTACAA 764  
Db 152810 GCCTCTGATCATCTGGCTGGTTCGCGCCACGACACGCGGCTGGAGAGCGGTACGA 152869  
QY 765 CCACGGGACGCAAAAGACGGAGATATCGTCTGCGAGAACATCTTAAACATTATGATAGG 824  
Db 152870 CCATACACGCAAAACAGACCGTGGCGAGCGGTGCTCTCGCGCGTATTTCA 152929  
QY 825 GAAGAAGGTGCTATCTGATGACGTACACTTGGAGAGAACTGGCGCTCTTTTAAACAA 884  
Db 152930 GAACGAGGTGCTCATCAACTACTCGACATCGCGCAACTGGGTCTATTTTACCAA 152989  
QY 885 CACGCGTGCACGATCTGAGTTCGACACCTTCTTAAAAATTTAGAGTATATCTTCC 944  
Db 152990 CAACGCGTTGCACGACCTTGACCTTTGGCGACCTGCTCAAGTACATGAGTTCGTTATCC 153049  
QY 945 CATTCTGACGAAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004  
Db 153050 GCTCAGGAGGAGCAGTGGCTTACGTGATCGTTCGCGCGCCCGCAGCTCCATCG 153109  
QY 1005 GATCTCTTTAGATAATATTGGCTGACCCGACTCTTTTGTGACATAGGAGTAAAGC 1064  
Db 153110 CAACCCCTCGTACGCCCGGAGATGACGCGGCTTTTCTCTCACAGTCTGAGTGCAGC 153169  
QY 1065 GTATCTAACAACTACACACATGCTTAACGCTAGGGGGGATACACGAGGACACGAGCT 1124  
Db 153170 TTACTTCAGGGGTGCACGACAGATCCGCTTGAGGAGCGCCACCGTCTTTGATCT 153229  
QY 1125 CTTTACCTGCTCCGTGGTGGAGGTTGTTGAAAGCCGTTTGGAGATACAAAGCGGGC 1184  
Db 153230 GCCCGTCTACTGGTGTGTAACACCGCGGTACCAAGAGCTCTGGGAGTGGCCGACCC 153289  
QY 1185 CGTCAACTCACCACCTCACTCCGTGACGAGTGGGTGACAAAACCTTTTAAAGTAA 1244  
Db 153290 GCTGGGCGATCGCCGCGGCGGTGAGCTCTGTTCCGCGAGACTTGGCGCGCATCAT 153349  
QY 1245 TAATTATTCGAGTTTGTGGACGAGGACATGTCATATGTCGCCACGGAATCCACTGAAC 1304  
Db 153350 TAACCTACTCGAGTTTGTGACACACACCTCTCCAGCGAGATCACCAAGAGGCGTGG 153409  
QY 1305 TTGACCCAGGTACCTTTTATACCAA 1331  
Db 153410 CCCCGCGCGGACGCTGTTGCCACCAA 153436

## RESULT 15

AAH74201  
ID AAH74201 standard; DNA; 124884 BP.  
XX  
AC AAH74201;  
XX  
DT 15-OCT-2001 (first entry)  
XX  
DE Nucleotide sequence of the Varicella virus Dumas strain.  
XX  
KW Dumas strain; pox vaccine; vaccine; poxvirus disease; chickenpox; ss.  
XX  
OS Varicella virus.  
XX  
PN WO200156600-A1.  
XX  
PD 09-AUG-2001.  
XX  
PF 31-JAN-2001; 2001WO-JP000678.  
XX  
PR 31-JAN-2000; 2000JP-00062734.  
XX  
PA (OSAU) UNIV OSAKA.  
XX  
PI Gomi Y, Sunamachi H, Takahashi M, Yamanishi K;  
XX  
WPI; 2001-488845/53.  
XX  
PT Method for checking absence of mutation at specific positions of  
PT varicella virus genome for quality control of attenuated live varicella  
PT vaccine.  
XX  
PS Claim 1; Page 61-105; 158pp; Japanese.  
XX  
CC The present sequence represents the nucleotide sequence of Varicella  
CC virus, Dumas strain. The specification describes a method for the quality  
CC control of an attenuated pox vaccine, characterised in that the DNA  
CC sequence of specific parts of the viral genome in a sample is determined  
CC and proved to be conserved rather than mutated in comparison with the  
CC same parts of a reference viral genome. The method is useful for quality  
CC control of vaccines for use in the prevention of poxvirus diseases such  
CC as chickenpox  
XX  
SQ Sequence 124884 BP; 33789 A; 29295 C; 28177 G; 33623 T; 0 U; 0 Other;  
Query Match 1.7%; Score 152.4; DB 5; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 2.7e-32;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 585 CAATATTATTGTAATCGACGAGCGCGGAACCCCTGCTCTTACATCTCACACCGTGT 644  
Db 96706 CAACGTTATTGTTATGATGAGCAGGATTCCTAGGCGGTCTATTTCTCAGCGCGTGT 96765  
QY 645 GTTCTTTTACTGTTTCTTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGCGGT 704  
Db 96766 TTACTGTTGGTGGCTTTTGAATGCTATATATCAAGGCCCTCAGTACATAAACGGTCA 96825  
QY 705 TCCGTGCAATAGTATGCTGCGTTGCCACGACGACAAACGCTTCCAGTCAACTTACAA 764  
Db 96826 ACCGCTCATAGTATGCTGCGTTGCCACCAACCTGACTCGTTAGATCTCATTTTCA 96885  
QY 765 CCACGGGACGCAAAAGACGAGATATCGTCTGGGAGAACATCTCTAAACATTATGATAGG 824  
Db 96886 ACATGACATGACGGTTTACACGTTACTCTAGTGAATAATATCTACGATATAATCTG 96945  
QY 825 GAAGAAGTCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 884  
Db 96946 CAATCAAACTCTGGCTCAATATCTAAACATCTCAATCACTCACTCACTCACTCA 97005  
QY 885 CAAAGCGTGCACCGATCTGCAAGTTCGGACACCTTCTAAAAATTTTAGAGTATATCTTCC 944  
Db 97006 CAAACGATGTCAGAGAGCAGATTTTGGAAATCTTTTAAAGCGCTTGAGTACGGCTACC 97065

Qy	945	CATTCTGACGAGTCATGAGTTACGTAGACAGATTGTGCTTCCTAAAGTAAGATTAT	1004
Db	97066	TATTACCGAAGCACATCGCGCTCTGTCGATACATTGTGTACCTGCTCCTATATTA	97125
Qy	1005	GGATCCTTTAGAATACATTGGCTGGACCGACTTTTGTACATAGCGAGGTAAGGC	1064
Db	97126	CAATCCTGCTAATCTTCCGGATGGACGCTCTGTATTCTGCATAAGGAGGTGAGCG	97185
Qy	1065	GTATCTAACAAACCTACACATGTCTAACGCTA	1098
Db	97186	GTATAGAGTAAGTTACACGGCATTTAAACTA	97219

Search completed: April 10, 2004, 22:00:17  
Job time : 2133 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 20:55:09 ; Search time 393 Seconds  
(without alignments)  
12811.882 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038  
Perfect score: 9073  
Sequence: 1 atggagagtcctcgatg.....ggggccacgagtcgcttga 9073

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTCUS COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfileseq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1002	11.0	35100	2	US-08-770-379-18
2	1002	11.0	35100	3	US-08-757-669A-18
3	1002	11.0	35100	4	US-09-230-371A-18
4	714	7.9	35100	2	US-08-770-379-19
5	714	7.9	35100	3	US-08-757-669A-19
6	714	7.9	35100	4	US-09-230-371A-19
7	154	1.7	125157	4	US-09-913-514-2
8	152.4	1.7	124884	4	US-09-661-596A-76
9	152.4	1.7	124884	4	US-09-913-514-1
C 10	137.6	1.5	1532	4	US-09-651-656-36
C 11	137.6	1.5	1532	4	US-09-650-855-36
C 12	134	1.5	723	4	US-09-489-039A-4593
C 13	125.8	1.4	7316	3	US-09-289-254-3
C 14	117.6	1.3	1830121	4	US-09-557-884-1
C 15	117.6	1.3	1830121	4	US-09-843-990A-1
C 16	111	1.2	4161	3	US-08-790-517-1
C 17	109.4	1.2	4161	3	US-08-790-517-1
C 18	94	1.0	738	4	US-09-252-991A-13238
C 19	91	1.0	119211	4	US-09-596-002-40
C 20	90.2	1.0	657	4	US-09-134-001C-2483
C 21	87.8	1.0	768	4	US-09-543-681A-3433
C 22	84.2	0.9	640681	4	US-08-790-988-1
C 23	84	0.9	8753	4	US-08-426-630-2
C 24	80.6	0.9	564	4	US-09-107-532A-353
C 25	74.8	0.8	729	4	US-09-328-352-3484
C 26	63.4	0.7	3662	4	US-09-221-017B-808
C 27	61	0.7	17846	4	US-08-956-171E-109

C 28	50.6	0.7	504	2	US-08-747-887-22
C 29	59.2	0.7	4403765	3	US-09-103-840A-2
C 30	59.2	0.7	4411529	3	US-09-103-840A-1
C 31	49.4	0.5	19124	2	US-08-487-826B-13
C 32	49.2	0.5	11049	4	US-10-204-708-23
C 33	46.6	0.5	5152	4	US-10-204-708-74
C 34	45.6	0.5	505	4	US-09-621-976-15639
C 35	44.2	0.5	6801	4	US-10-204-708-61
C 36	44.2	0.5	6865	4	US-10-204-708-20
C 37	42.8	0.5	5852	1	US-07-867-106-2
C 38	42.8	0.5	9347	4	US-10-204-708-35
C 39	42.4	0.5	1230025	4	US-09-198-452A-1
C 40	42	0.5	6113	4	US-10-204-708-14
C 41	41.8	0.5	474	4	US-09-621-976-18033
C 42	41.8	0.5	53332	4	US-09-801-861-3
C 43	41.8	0.5	640681	4	US-09-780-988-1
C 44	41.6	0.5	1001	4	US-09-671-317-141
C 45	41.2	0.5	3740	4	US-09-489-847-14

ALIGNMENTS

RESULT 1  
US-08-770-379-18  
; Sequence 18, Application US/08770379  
; Patent No. 5849564  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Echenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
; VIRUS: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/770,379  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 52342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35100 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-770-379-18

Query Match 11.0%; Score 1002; DB 2; Length 35100;  
Best Local Similarity 54.5%; Pred. No. 5.1e-292;  
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;  
QY 1 ATGGAGAGTTCGTCGGATGGACCAACACGTCGACCAATCCGGGGTTCATCTTGAAC 60

Db 29792 ATGGACAGCTCGGAGGGTGCATGATATGGACGAGCCAGCCCGGGTTCATCTCAAC 29851  
 Qy 61 ATGACGTCCGATCCCAAGTCAAGGGTGTGCTGGATCAGTCAGTCGCTGCAATATA 120  
 Db 29852 ATGACATCGATCCCAAGTCAAGGGTGTGCTGGATCAGTCAGTCGCTGCAATATA 29911  
 Qy 121 ACTACAGCCCAACCGGAATGGTGTGATGACGACCTGGCTTCGATCCGGCTGAAGATCC 180  
 Db 29912 ACTACCTCGCCCGGAGATGGCTGGATGACCTAGAGTTCGATCCACTGGAAGACGA 29971  
 Qy 181 GGSCCGTCTTGCCTTACCTGTTTATCTAATTCAGGGAATGCTGGTGGTGGAAAGT 240  
 Db 29972 GGCCCTCTTCTGCGCTTTTCGGAATGCTGATTAACGGGAGCTGAGGAGCGGGAAAGC 30031  
 Qy 241 ACCAGCATATCGCCCTGTACCAAAATTTAAATCTGCTGATCAAGGGGCGGACCAACATA 300  
 Db 30032 ACCAGGATATCGCCCTACATCAGAACTCTCACTGCTTAATACGGGGCTACAGTGGA 30091  
 Qy 301 GCGGCACAGAACTATCGCTGCTGCTTAAGAGTTCCTGCTCCAGATCTTCAGCGCTTTT 360  
 Db 30092 GCGGCACAGAACTTTTCAGGGCTTTAAAGTCTTACTGTCTCCACTATATACCGCTTC 30151  
 Qy 361 GGCTTTAAGCGGACACATCAATATAGCCGTCAAGAAAGCTCATAGACGGAGCCGTA 420  
 Db 30152 GGATTCAGAGGACAGACATTAATATCTGCCAGAGGAAAGTCCCAAGTAACTCAGTCC 30211  
 Qy 421 TCCATAGACAAATTCAGCAACAGAGCTATCGAGTATTCGCGGTATAGTGACAT 480  
 Db 30212 TCCATCGACAACTCCAGAGATACAGAGTGGCTAGTACTGCGCACTATATACCGCTTC 30271  
 Qy 481 ATGAAGAGGTATGGCGAAACAAACCAATGCGATGTACGGGACTATCAACCGCAAT 540  
 Db 30272 ATTCGAGAAATTTATGCGAAGAAACAAAGGGGAGTATAGTCCCTCTCTCAAGCGCT 30331  
 Qy 541 TTGAAACCTCTCGAGATGACCGGACCGTGTATGAGTTCCTCAATATTTATGTAATC 600  
 Db 30332 TTGAGACTCTTTCGCTATGGTGGAGCAATTTGCGACGAGTAACTATTCGTAAT 30391  
 Qy 601 GACGAGCGGAACTCTGCTTCTTACATATCAACCGCTGCTGCTTTTACTGCTTC 660  
 Db 30392 GACGAGCTGGAACCTCTGCTCCATATTTTACCGCGGTGCTGCTTCTTATGCTTT 30451  
 Qy 661 CTAAACAGCTGCTAAATACCCCTCTTTACCGCAAGGGGGGGTTCGCTGCTATGATGC 720  
 Db 30452 TACAACAGTGGCTGGACACCCCGCTATACGAATTTGTCGCTGCTTCTGCTATGCTGC 30511  
 Qy 721 GTGCTTCGCCAACGACGAAACCGCTTCCAGTCAACTTACACCAAGCGGACGCAAAAG 780  
 Db 30512 GTGGGCTCTCCACCCAGAGGAGCGCTTTCAGTCCGTCTTCAACCAACGACGAGAGA 30571  
 Qy 781 ACGGAGATATCGCTGCGAGAACATCTTAATTCATGATAGGAGAGAGTCTGATCT 840  
 Db 30572 AACGAGATATCGCTGCTGATATGCTGCTCACTTCTTATGGGAAACGCTGAGGTGCA 30631  
 Qy 841 GAGTACGTACACTGGAGAGAACTGGCGCTGTTTATTAACAAACAAAGCGCTGACCGAT 900  
 Db 30632 GATTATATAGGCTGGAAGAGAAATGGGCGCTTATTAACAAATAGCGCTGTACGAT 30691  
 Qy 901 CTGAGTTCGACACTCTTAATAATTTAGAGTATATCTTCCATTCCTGACGAGTCT 960  
 Db 30692 CCCCAGTTGGTCACTTCTGAGACCTTGAATATATCTATGACATATCAACAGAGTTA 30751  
 Qy 961 ATGAGTTAGTACAGATTTGTGCTTCTTAAAGTAAAGTATGATTCCTTTAGAAATAC 1020  
 Db 30752 ATGGACTATATAGATAGTGTGCTTCCGAAGAGTAAAGTTCCTGGACCCGCTCGAGTAT 30811  
 Qy 1021 ATTTGCTGACCCGACTCTTTTGTACATAGCGGCTTAAGCGGTATCAACAAACCTTA 1080  
 Db 30812 GCAGGTGGAAGAAGCTCTTATCTCACACGAGGAGTGAAGTCTTTTCGGCAACGCTG 30871  
 Qy 1081 CACACATGCTAAACGCTAGGGGCGATACCAAGGACAGAGCTCTTATCTGTCCGCTG 1140

Db 30872 CACACCTGCTGCTCGAGTAATAAGATGCTGTGTCACAAAGCTTTTACCTGCCAGTG 30931  
 Qy 1141 GTGTGCGAGGTGTTTGTGAAGCCGTTTGAAGAAATACAGCGGGCCCTCAACCTCACCAAC 1200  
 Db 30932 GTCTGTGAGGTGTTTACAGGCCATTTGAGAGTACAAACGGGCGTAGGCTCACACAC 30991  
 Qy 1201 CTCACCGTACCGAGTGGGTGACAAACCTCTTTAAGTTAAGTAATTTATTCGAGTTT 1260  
 Db 30992 ATGACTCCCATAGAAATGGTAAACAAAATCTTTTACGGCTAAGTAACACTACTGCACTT 31051  
 Qy 1261 GTGGACAGGACATGCTCATAGTCGCAAGAAATCCACTGAACTGCTGACCCAGTTTACC 1320  
 Db 31052 GCTGATCAGGACATGGCTGTGTTGGACCTTATATCACAGACGCTCACACAGATCACC 31111  
 Qy 1321 TTATPACCAAGTTTGTPTAAAAACAGCACGCTGCTTTTAAATGAAAAACAAAAATGC 1380  
 Db 31112 TTGCGCACTAAATTTGTCAAAAACAGCTATGCTACCTTTACTGGAAGACCAAAAAATGT 31171  
 Qy 1381 ATATGCTGGTTCAGGCTACGCTATCTTTGAGTTCAAAAGATCTTAGACAGGACTCTTC 1440  
 Db 31172 ATATCGGGTTCACGGTCTATCAAAAGATTTCAAGTCCATCTTAGACGGGAGCTATTT 31231  
 Qy 1441 GTGGAGACGCTATCGCAAGACCTCCGGAATACGTATPACGCTTTCTTAAACACGCTACTG 1500  
 Db 31232 ATCGAAAGTCAATCGCACATAACCCGCTTATGTGTACAGTTTCTTAGTACCTGCTA 31291  
 Qy 1501 TACACGCCATGTAATCTGTTTACGCTGAGCGTACGCTAGTACACAGGAAATATCTG 1560  
 Db 31292 TATATGCAATGTAATCTATTTTACGCAAGGGTGAAGCAGGGGCAATGAAGATTCCTC 31351  
 Qy 1561 CAAAGCTTAAATTTGCAACCCCTCCCGCGCTCTGGCAACCGGGCGGTGAGACCTTCAA 1620  
 Db 31352 AGGGACCTCAGGAACTGCGGCTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 31411  
 Qy 1621 ACGGTTCTGAGAGTTTAAACCTGGAAGACGACATCTTACACGCTGCTGCTCCCGCG 1680  
 Db 31412 GTTCTGGGGAGAG-----GGGGACACAGATGCTTCTACCTCACCGCAGCTCCCA 31465  
 Qy 1681 CCACCGCGGTATCACCTCCCTCCAGGTTTGTGTGACACAGCTATCTGCGCTTAAAGGAC 1740  
 Db 31466 CCATCCCGCAACCCAGCGGCTCTTCAACACTGGTGGCTTATTTACTCCGGGGCAAGAA 31525  
 Qy 1741 GTGTTGCGCTCCAGAAATAAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
 Db 31526 CTATTTGCAACGCTGGCGCTGGACGCGGACACTTTGGTGAAGTTCTTCCCTACTCC 31585  
 Qy 1801 ACGTTTTCGCGTTTACCGTTTAACTGCTGAAGGACGAGTGTGATTTGTCTCCCT 1860  
 Db 31586 GATTTTCAACGTTTACCGTGAACATCTGCTGCGAGATGGCTGGACTTTGTGCTCACT 31645  
 Qy 1861 TCAGAACGCTTCAACGGGCTGTTGGCTTTGCTATCGACCGTTCGATCGTATTAATTAAG 1920  
 Db 31646 TCCCGCGCTCCACGGCTAGTGGCATACGCTACACTATAGACCTATATATATCCAG 31705  
 Qy 1921 GGSTACACGTTTTCACCGGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980  
 Db 31706 GGATATACGTTTCTCCAGTGAGATTCGCGCTCCAGGAGGACAGCGCTCAGCGAGGAC 31765  
 Qy 1981 CTCAGGAGAGATGCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
 Db 31766 CTGCGCAAGAAATGCTTCTCATAGTTGCTCAGGATCTATCGGGTTCATTTGCTGCTG 31825  
 Qy 2041 GAGAAATACATAACCAAAATTTGACCGAAACCATGAGGAGCGGAGCGGTTTCCAAAGTGTGC 2100  
 Db 31826 GAAATAACGCTACCAAGATGACAGAGCCCTCGAAGGTGGCGACGCTGTTTAAACATATGT 31885  
 Qy 2101 TGTGCGGGGACTATGCGGTGAGTCAAAATTTAGCCATGACCATCTGTAAGGACAGAGGA 2160  
 Db 31886 TGTGCGGGGACTAGGATATGAGTTCTATCTGCTGCTATGACCATAGTGAAGGACAGGGG 31945  
 Qy 2161 ATGCTGTTGAGGAGGAGTACCGGTAGTATTTGGTCCCAAGAACGCTCCAGACAAACCCAC 2220  
 Db 31946 GTTTCACTAAGTGGGTGGCCATATCGTTCCGCAACCAACCGCAATATACAGAGCAGTCTA 32005









3379 GTCTTGTCTCTTTTCCCTAGTGGCCCGGAAATCGGAGCATCTCTGGAGTGGCCGCG 3438  
3386 TCACGGGGGCGTCTCTCGAGCGCTCTTCAATTTGGAGGATCTCTCTTCATCGTGTG 3445  
3439 GGTCTCTTGAAGCGGTCCGATAGAAC-----ATAGCATGCGC 3476  
3446 TGTAGACAGAGGCTCTCAAAACATCGCCATCGCCCTTGTACGGGTTGACCGCTAGGGC 3505  
3477 GAAACGCTTCACTGTAAACGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGG 3536  
3506 GGAATTTACAAAGACACAGGATTTATGCTTCTCAACAGAGCCCGCAGTCCACAG 3565  
3537 CACCCCTCGATCTGAATAGATAGTCAATTCGCAAGTTTAAATGTTTGAAGCTAGAA 3596  
3566 TCTACCGCGGTGGCGAGTCAATAGTGTGGCTAGGTTAAGTATTTACAGCCCTGGA 3625  
3597 ACTGCGCCGAGGAGAACCGGTTGCGGTGGCGGTAGCCAGGGGAGACGGATGGCGG 3656  
3626 ACCGAGGCATCGAGATGTCGCCACCAAGAGAG--GCCAGCGAGATGATGCTGG 3682  
3657 CTTGAGGACGAGATGCTCTGCTGCTTTATCAGGCTGCGCTTGTCTATAGCCCTTCTC 3716  
3683 CCGTAAGCACAGGTTTCTGCTGCTTTATGAGCGGAGTTCTGTCAATGSCCTTGGCC 3742  
3717 CCACAGCATATAAAGCAGCTCTGAAGCTTGGCAGAAAGCAGCTTATATGATGCTGG 3776  
3743 CCACAGGAGAAACCGCAATGTTCTAATTTGAGGATATGCTACTGATGATGAATCG 3802  
3777 TGAACCAATCCAGCCAGGTTGGGCTGTGAGCCCGCTTCCCTCTTCCACCGTCAAA 3836  
3803 TGAACCAATCCAGCCAGGTTGCTCTGTGAGCGCGCCCTCCCTCTTCCACCGTCAAA 3862  
3837 TGGTGTAAAAAGCAGAACTCCCGCTTTTGGCCCAACAGTCTAAGCAGCCGCTGAGAAAGG 3896  
3863 CTGTGTTTATGAGCAACACACCTGCGAGCCAGCTGTCTGAGCAACCCGTTGGGAGGAG 3922  
3897 CGGTGACCCAGTACAGTATTTGGCTATCTCTTTAAATATTTTGAAGCTTGGAGGTA 3956  
3923 TACTGAAATTTGGGACGGAAGCTCTAGCTCTTAAGATGCTCTCAAACTGGGTGGA 3982  
3957 TACTGTAGTCCGAGCAACGCTAAAGGCTAGTCCCGTGGCTTGAACCGGATGTAAGGAT 4016  
3983 CTGTGATTTGGATCCACACTTAACGCGAGCCAGTGTGCTTGGCTTGGTACGGGT 34042  
4017 CTGGCCAAAGATACAGACTTTAATATCTCTCGGTTCGCAACAATAAGACCAATAA 4076  
34043 CTGGCCCTAAGATCAACACTTTAATATCTCTGATCGCAGCAGTGGGACCAACATCA 34102  
4077 TCTATCTATCGTGGATAAACACGCTTACGTTGCTCATATCCATTAACGCGCTTCA 4136  
34103 GCTTGTCTGTGGGGATACATGCTGGTGTAGCTTAAGTTCCGATCTGTCTGAGCA 34162  
4137 GAGCCGCGAGCTCTGTTTAAATTTGAGAAAGATTTAAATTTCAAGCA 4188  
34163 GCGAGAGCAGTTCTGTTTTCAGAAATGATGAGGCTCAGAAAGGAATCCA 34214

## RESULT 3

US-09-230-371A-18  
; Sequence 18, Application US/09230371A  
; Patent No. 6348586  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A  
; APPLICANT: Russo, James J  
; APPLICANT: Edelman, Isidore S  
; APPLICANT: Moore, Patrick S  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
; FILE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 45185-G-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,371A  
; CURRENT FILING DATE: 1999-11-17

; PRIOR APPLICATION NUMBER: PCT/US97/13346  
; PRIOR FILING DATE: 1997-07-22  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 35100  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-230-371A-18

Query Match 11.0%; Score 1002; DB 4; Length 35100;

Best Local Similarity 54.5%; Pred. No. 5.1e-292; Indels 253; Gaps 10;

Matches 2414; Conservative 0; Mismatches 1765;

QY 1 ATGAGAGTTCGTCGGATGACCAACACACGTCGAAACCAATCCGGGTTTCATCTGAAC 60

DB 29792 ATGGACAGTCGGAAGGTCACCTGATATGACGAGCCAGCCCGGTTTCATCTCAAC 29851

QY 61 ATGACGTCGGATGCCAAAGTCAGGGTGTCTGATCAGCTCAGTCCGCTGTCATAATA 120

DB 29852 ATGACATCGGATGCCAAAGTCAGGAGCGTAGTGAACAGATTGACAGTTGTCAAATATC 29911

QY 121 ACTACAGCCCAACCGGAATGGTTGGTACGACCTGCGCTTCGATCCGGCTGAAGATCC 180

DB 29912 ACTACCTGCGCCCGGAGATGGGCTGTATGACCTAGATTCGATCCACTGGAAGAGAA 29971

QY 181 GGGCCGTTCTTGGCGTTTACCGTTTATCTAATTAAGGGAAGTCTGCTGTGTTGGGAAAGT 240

DB 29972 GGGCCCTTCTTGGCGTTTTCGGCATACGTAATAAGGGAAGTCTGAGAGCGGGGAAAGC 30031

QY 241 ACCAGCATATCGGCGCTGTACCAAAATTTAACTGCTGATCAAGCGGCGGACCAACATA 300

DB 30032 ACCAGCGTATCGGCGCTGTACCAAAATTTAACTGCTGATCAAGCGGCGGCTACAGTGGTA 30091

QY 301 GCGGCACAGAACTATCGCGTCGCTAAAGACGTTCTGCCAGCATCTTCAGCGCTTTT 360

DB 30092 GCGGCACAGATCTTTCAGGCGCTTAAAGTCTTCTGCTCCACTATATACCGCGCTTC 30151

QY 361 GCGTTTAAAGCGGACACATCAATATAGCCGTGAGAAAGCTCATCAGACGGAGCGGTA 420

DB 30152 GGAATTAAGAGACACATTAATATCTGCGAGAGAAAGTGCACCAAGTAACATCAAGTCC 30211

QY 421 TCCATAGAGCAAAATTCAGCAACAGAGCTATCGAAGTATTTGGCGGTTATAGTGGACAT 480

DB 30212 TCCATCGAGCAACTCCAGAGATACAGCTGTGCTAGTGTGCGCACTGTACCGATATT 30271

QY 481 ATGAAGAGGTTATGGCGAATAAACCACATGGCATGTACGGGACTATATCCACGCGAAT 540

DB 30272 ATTCAGAAATTTATGCGCAAGAAACAAAGAGGGGAGTATAGTCCCTCTCTCAAGCGCT 30331

QY 541 TTTGAAACCTCTCGAGAAATGACCGGACCGTGTGTTATGGACTTCCCAATATATTGTAATC 600

DB 30332 TCCAGCTCTTTGGCGTATGGGTGGAGCCCAATTTGAGACGAGTAACATTTCTGTGATA 30391

QY 601 GACGAGCGGAAACCTGTCTCTTATACATCACTACCAACCTGTGTTCTTTTCTGTTTC 660

DB 30392 GACGAGCTGGAACCTCTCTGTCCTCATATTTTGAACCGCGTGTGTTCTTCTATTGGTTT 30451

QY 661 CTAAACAGCTGGCTAAATATACCCCTCTTTTACCGCCAGGGGCGGTTCGCTCATAGTATGC 720

DB 30452 TACACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTGCTGCTAGTCTGC 30511

QY 721 GTCGGTTCGCAACGCGAGACAAACGCTTCCAGTCACTTACCAACGAGGAGCGGAAAG 780

DB 30512 GTGGGTTCTCCACCCAGAGCGGCTTTTCAAGCGGCTCTTCAACCAACGAGGAGAGA 30571

QY 781 ACGGAGATATCGTGTGCGAGAACATCTAACTTATGATAGGGAAGAGGTCGTATCT 840

DB 30572 AACGAGATATCTGCTGTGATATGCTCACTTCTTATTTGGGAAACGAGGTTGCA 30631

QY 841 GAGTACGTACACTTTGGAGAGAACTGGGCGCTGTTTATAAACAACAGCGCTCCCGAT 900

DB 30632 GATTATATTAGCTGGACGAGAAATTTGGGCGCTTATTTATAAACAATAAGCGCTGTACGGAT 30691

QY 901 CTCGAGTTCGGACACACCTTCTAATAATTTTAGAGTAAATCTTCCATTCTCGACGAGTC 960  
 Db 30692 CCCAGTTTGGTTCACCTTGTGAAGACCTTAGAATAATACTAGACATATCACAGAGTTA 30751  
 QY 961 ATGAGTTACGTAGACAGATTGTGCTTCTTAAAGTAAGATTATGGATCCTTTAGAAATAC 1020  
 Db 30752 ATGACATATATAGATAGTTTGTGTTCCGAAGTAAGATTCTGACACCGCTCGAGTAT 30811  
 QY 1021 ATTGGCTGGACCGACCTTTTGTTCACATAGCGAGTAAAGCGTATCTACAAACCTA 1080  
 Db 30812 GCAGGTGGACAGACTTTCTATCTCACACGAGGTGAAGTCTTTCTGGCAACGCTG 30871  
 QY 1081 CACACATGTCTAAGCTAGGGGGCGATPACGAGGACACGAAAGCTCTTTACCTGTCCCGTG 1140  
 Db 30872 CACACCTGCTGTGCGAGTAAAGGATGCTGTCTCCACAAAGCTTTTCACTGCCAGTG 30931  
 QY 1141 GTGTGGAGGTGTTGTGAAGCGTTTGAAGATACAGCGGCGCTCAACCTCACCAAC 1200  
 Db 30932 GTCTGTGAGTGTTTACAGAGCCATTTGAGGAGTCAAAAGCGGCGTAGCGCTTCACAC 30991  
 QY 1201 CTCACGCTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTATTCGCAGTTT 1260  
 Db 30992 ATGACTCCCATAGATGGGTAAACAAAATCTTTTCAAGCTAAGTAACTACTGCGAGTTT 31051  
 QY 1261 GTGGACAGGACATGTCATAGTCGCGACCGAATCCATGAACTGTCGACCCAGGTTACG 1320  
 Db 31052 GCTGATCAGGACATGGCTGTGGTGGACCTATATCACAGACGCGTCCACAGATCACG 31111  
 QY 1321 TTTTATCACCAAGTTTGTAAACAGCCAGCTGTCTTTAATCGAAAACCAAAAATATG 1380  
 Db 31112 TTGCCACTAAATTTGTCAAAACAGCTATGTCACCTTACTGGAAGACCAAAAATGT 31171  
 QY 1381 ATATGTGGGTTTCAGGTTAGCTACTTTGAGTTCAAAAGATCTCTAGACAGCGAATCTTTC 1440  
 Db 31172 ATATGGGTTTCACGGGTATACCAAAAGATTCAAGTCCATCTAGACGGGAGCTATT 31231  
 QY 1441 GTGGAGAGCGATTGCGAAGCCGTCGGAATAGTATACGGTTTCTTAACACGCTACTG 1500  
 Db 31232 ATCGAAAGTCAATCGCAGCAATAACCCGCTTATGTGTACAGTTTCTTGTAGTCCCTGCTA 31291  
 QY 1501 TACAACGCCATGTACTGTTTACGCGTACGGGTGACTAGGTCAACAGAGAAATATCTG 1560  
 Db 31292 TATTAATGCCATGTACTCATTTTACGCGCAGGGGTGAAGCAGGGCATGAAGATTCCTC 31351  
 QY 1561 CAAGACTTAAATTTGCAACCTCCCGGCGCTCTGGCAACCGGGCGGTGTAGACCTTCAA 1620  
 Db 31352 AGGACCTCAGGGAAGTGCCTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 31411  
 QY 1621 ACGGTTCGTGAAGAGTTAAACCTGGAAGACGACATCTTACACAGTGTGCACTCCCGCG 1680  
 Db 31412 GTTCTGGGCGAGGAG-----GGGGACACAGATGCTTCTACCTCCGCGCAGCTCCCA 31465  
 QY 1681 CCACCCGCGGTATCACCTCCCTCCAGGTTTGGTGCACAGTACTGCGCCCTAAAGGAC 1740  
 Db 31466 CCATCCCCACCCACCGGCTCTTCCAACTGTGGCTTATTAATCCGGGGCCAAAGAA 31525  
 QY 1741 GTGTTGCGCTCAGAAATAAGGTGCGTGTGCTGGTTTGGCGGGAGTTTGAGAGGAA 1800  
 Db 31526 CTATTCGCAACAGGCTGGCCCTGGACGCGGACACTTTGGTACGAGTTTCTCCACTCC 31585  
 QY 1801 ACGTTTTCGGGTTTACGGTTTAAATGCTGTCTGAAGGACGAGTTGACTTTGTCTCCCT 1860  
 Db 31586 GATTTTTCAGGTTTACGGTGAACATCGTGTGCGAGATGCGTGGAGTTTGTGTCACCT 31645  
 QY 1861 TCAGAACCTCTCAACGGGCTGTGGCGTTTGCATCGACCGTTGAATCGTATATAAATTAG 1920  
 Db 31646 TCCCCCGGCTCCAGGCTCTAGTGGCATACCATCCACTATAGACACTATATAATCCAG 31705  
 QY 1921 GGGTACAGTTTATACCGGTAGCGTTGCTGCTGCTCAGGTTTGGCACTCAGTATGAC 1980  
 Db 31706 GATATACGTTCTCCAGTGAATTCGGCCGTCCAGGAGACAGCGCTCAGCGAGGAC 31765

QY 1981 CTCAGGAAGAAGATGCCCTCCCTCGTGTGAGGACCTCTAGCGGTTTTTATCGCGTGCTTA 2040  
 Db 31766 CTGCGCAGAAAGATGCCCTCCATAGTTGTCTCAGGACTCATCGGGTTTCAATTTGCGCTGCTG 31825  
 QY 2041 GAGATAACATAACCAATTTGACCGAAACCATGAGGACGAGCGTGTTCCTCAAGTGTGC 2100  
 Db 31826 GAAATTAACGTGCAACAGATGACAGAGACCTCTCGAAGTGGCGAGTGTGTAAATATGT 31885  
 QY 2101 TGTGCGGGGACTATGGGGTTCAGTCAAAATTTAGCCATGACCATCTGTAAGGACAGGGA 2160  
 Db 31886 TGTGCGGGGACTACGGTATCAGTTCTAACTCTGGCTATGACCATAGTAGTAAGGACAGGG 31945  
 QY 2161 ATGTCTGTGGAGCAGTAGCGGTAGTATTTGGGTCCCAAGAGACGTCACAGAGCCAC 2220  
 Db 31946 GTTTCACTAAGTAGGGTGGCCATATCGTGTGGCAACCAACCCCATATCAGAGCCAGTCTA 32005  
 QY 2221 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAAATTTATTTGGTCATGGACAGCAACCCC 2280  
 Db 32006 GTGTATGTGGGTGTATCCAGGGCCATCGACGCTCGTTACCTGGTAAATGACAGTAATCCC 32065  
 QY 2281 CTTAAACCCCTCTCAGAGAACAGTGTGATTAACCTCCGCGCAAGCATATAGTCCGCGCC 2340  
 Db 32066 CTTAAGCTAATGGACCGCGGTGACGCCAGTCCCATCTCTCAAAGTACATCATCAAGCC 32125  
 QY 2341 CTCACACACCCAAACACAAACCTCATCTACTAA----- 2373  
 Db 32126 CTATGCAACCCCAAGACTCTCTGATCTACTGACCCGTAACCCCTCTCTTAGACACTGAT 32185  
 QY 2374 -----ATNAATAAAACAAAGCAAAATGGAAAAATTTG 2406  
 Db 32186 GTGTTTGGGAATAAAGCATGAGATTGACACCTATAATGGTCTGTATTTGACACCAATCTT 32245  
 QY 2407 TTTTATTTACGTCCAAACACCGCCACGGATAGTTGTCAATTTCCACACACCGGGGGGTGG 2466  
 Db 32246 TTTATTTACAGTCCAGCCGCGAGTTATATGCAACGTTTCCACACAGGGGTGGCGTGG 32305  
 QY 2467 CGGCATATGTTGACGACAGAAACGCGCGGCGCTCTGGGTAGAAATGAGCCCC----- 2521  
 Db 32306 AGGCCAAGATGCGGGTGGTGGTGCCTGACCTGGACCCCGGTAGTTGTGTTCTCGATG 32365  
 QY 2522 -----GCCCCCGTGTGAGGTCTACAGTCTCTTAGGGCCCTGGTT 2564  
 Db 32366 AAATCGAGTGGCGGAAGTACTGGGAGATTGGTGGAGGTGACCCCTTTGTGCTCGAG 32425  
 QY 2565 GTACCAACATAAATCAGCTGTCAACGTTGACGTCCTCGTGTGACGCGCGCTCTTTG 2624  
 Db 32426 GAGACAGATCAGCTCAGCGCGGACGAGGCTCTCTGCTGTGTCTCTCCCGAGGATA 32485  
 QY 2625 CAGCGGCGACGAAACCGCGCTTGTAAAGTATCT-----GCAT 2664  
 Db 32486 TAAATATCAGGACGCCACTGCTTTGCGGCTTAAAGTTGGTGTCTCTGGACGCGACA 32545  
 QY 2665 CGTGTGCTGTCTCAAGATGAACACAGGTCAATCAATCAATCAATGAGTGTCTCTGGCGCCCTCAT 2724  
 Db 32546 CATCTCGCTACAGAGGAGCGGTAGACTGCTTTTGGCTTTCTGGCCACGTCATGA 32605  
 QY 2725 CGCGCTGTGATTTGTCCTCATGTTCAACAGAGCCTCGGTGTCCTCACTTGTGCTCATCTG 2784  
 Db 32606 GCCCGATTCTCTGACTCAATACTTCCCTTGTCTCTCTCGCTCTCTCTCGGACGAGGTTG 32665  
 QY 2785 AAATTAATGTCTCTGCTGCTGCTGAGTGTGATCCAGCAACACAGGTTCTTCTTCTATCGG 2844  
 Db 32666 GTGTGTGGGAAAAATGGCGCGCTCGTAAACCGGCTCTATTGTTACGTCGAGAGT 32725  
 QY 2845 AAGACCATCCGCGCTCTGCGGTGATCTAGATATATCACTGACAGGCGGCGAGGCCCC 2904  
 Db 32726 TGGAACTGTCTCGCTATCAGAGTCCGATGTCAAGTCCAGTCCGCTGCGGTGGCGCGC 32785  
 QY 2905 G----- 2905  
 Db 32786 GCAGGGGCGCCACGAGGCGCTTTCATCAGGGTGTGTGTATGGTGAACCTTTGTGTTCAG 32845  
 QY 2906 -----TGGCGAAAAGCCCTCAGCTTCTTCGCGCACTTTTCATCAGTGT 2946



Db 549 GTGGTAGAGCAGCAGAGCTTCAACACGATGABATCAAGCTAACCTCCCGCGGAATC 608  
Qy 5631 CCGGCTCAAAACCGAATCTCTTAGGAACGACATCAAGCGTTACAGCAAAATTCGA 5690  
Db 609 CACATCCACAAAGGGGTATTGCTCGGGTCTCTGTATTAGGCTCGGAATAGAAAATCTAGA 668  
Qy 5691 AACCGAAACCTAGTTCTGAGCACCACATCTGGTTATAAATCTGGGTTAAGAGACAGAGCGC 5750  
Db 669 AAAAGACACTGACCCACACAGAGAGAACCTGGGCTCTTGCMAAGTTGATGAGCCCGCAGA 728  
Qy 5751 CACACACTGCTGCTGTGAGGANTCAGCTTAGGTGGCTGCGAGCTTCTGACAGGCCCG 5810  
Db 729 AAGAAATGTTCTCCCGTGGGACAAAGAGCTTGGGGGGGACAGATGGGCTACAGTGGGT 788  
Qy 5811 TAGCTGCCGTGGGGGCTCGGCACCTTTGACAGCGCAGCCAGCTTCCAGAGCAGGCA 5870  
Db 789 GAATTTCTTACACCGTATACATTGTTGGCACCACAGGCTGTTCCAGTATCAGCAT 848  
Qy 5871 GAAGCGCTCTTTGCACTCGCGCATATCTCGGCCATAGAGANTCGCTGGGTGATCTTT 5930  
Db 849 AAATCTATCTTTGCACTCATCCAGATCAAAAGTCAATGTCAGATGCTGTGCTGGCAATTT 908  
Qy 5931 TAGGTGCATATATATCTGTTGGTAAACAACTCACTCCTCTGTAAGAGGGAATAGGTC 5990  
Db 909 TGCCTGCTATCTTCTGCTGCCACATATTTAAACATCTCTAATCTGGAAGTAGATT 968  
Qy 5991 CGTTCTTCTATCTGTTCTCTGAATCTATAGTTACATCTTTAAGTATAAACCAAGAGCG 6050  
Db 969 CAGCTGTGGTTGAGCCCCCGGGAAGCAGCGGTATGCTTCAGGACCAACAGGAGCG 1028  
Qy 6051 TAAAAACCCAGGATGTCAGACCTTTTAAAAATGCAAAAGCGACCGATGGCTCTATC 6110  
Db 1029 TAAGAACCCCGGGTGTCCGGCTCCGGAACAGACCTCTGAGATACCGTCCGGTCTTGAC 1088  
Qy 6111 CGCGTTTCCATGCTTCAAGTGTACAGTGTATCCAGTGTAGACTCTGTCTCCCT 6170  
Db 1089 GAAACCGAGTGTGTACGATGCCAATCTGTGCGCTCCAGCTCTCAAAATTTTCATC 1148  
Qy 6171 GTTAAACCCCTTAACCTGGAATGAACAGCCATGACACCGTGAACCTTCTGTAGTCTCT 6230  
Db 1149 TCCAAATACCCGGAATTTGGGATACACACCTCCATGTTCACTCACTGATGAGGTCTC 1208  
Qy 6231 AAACCTAACCCGAAACAGAGCTAAATACCAATGCTGTCACCCCTACCCCAAGCAGCG 6290  
Db 1209 CCCACCCAAACCCCATAGGACCCAGCTACAGCTTATCTCCCAATAATACAGGAGCTA 1268  
Qy 6291 CCCCGTACTATT-----AGACAGGGGTGAGTAAACACCGCTATCCT 6330  
Db 1269 CCGGCACTCATTAAGCCCGCCAGAAACCACTAGCTGGGTGGCAATGACACGTCCTCT 1328  
Qy 6331 TTAATAACCCATACGTGGAGTTGTAGGTAAACAGCT-CTGTATATTTCAGACGCTG 6389  
Db 1329 TTAATAAGTCAACCTTACTCCGCAAGGGGTAGTCTGTGTGAGAACTACTGTCCAGCAGC 1388  
Qy 6390 TCATAAATGGATACCAAGCAATCAGGTAATAAATTTTATTTTCAAGATTGAGCA 6449  
Db 1389 CACAAATATGGCAAGTACAAAGGTAAAGATCGAC--TTTTTATTGTATGTAACA 1446  
Qy 6450 GGGCGTATTACAGTGTGTAGTGGAGCATATCTGCGCAATGGGAATGGCTGATGAAT 6509  
Db 1447 ATGCGTATTACAATGTTAGTGGAGCAGAGTTGCGCAAGCTCTACGTCCGAACAGT 1506  
Qy 6510 CCACATTAAGTGTCCGATCAGGTGCTCTGTGACTTTTACTAGTGGCATGAACCTTAAT 6569  
Db 1507 CGGGTGTGAGGCTCTTTATTAAAGTTTCCGTTACTTTCACAAAGCGCGGAACCTAGGT 1566  
Qy 6570 TTGGCGTGTTTAGTGTGATCAACCAAAACCTGCTGGCTCGCCTTTTGAAGCTCTA 6629  
Db 1567 TGGTCTGTACAGTCTGTACAGGCAAAAGAGATCGGGCGTGTCTTTTTCAGGAGAGTTA 1626  
Qy 6630 TGCAGCTTTAATCGTGTCCATAGCTGTCTTTTAAATGCAACTCTGTGATCATCTTTA 6689

Db 1627 GGGACGTGCTGATTATATGTGACAAAGCTTCTGCTCGTAAATGCAACCGCTGCTAGTAA 1686  
Qy 6690 TGACAGTGGTCCAAAAAACAAGATTTAAAAACAGCTTTAAAACTGCGTCTCTGCTAGT 6749  
Db 1687 CGACAGCTGTCAAAAAACAAGGTTCAGCTGCAAGTTAAATCTGTATCTCTGAAAGT 1746  
Qy 6750 CTTGCTATATAACCTCTCAACAGAAAAAATTTTAAACCAACTCGCCAGCTACTGAA 6809  
Db 1747 CTTGCTAAATGACAGTTTCTTACCAAGAAAAAATTTTAAACCGCTGGCCATCCACTGAA 1806  
Qy 6810 ACCATGTGTAGGACAGGTCTGTGAAAGTGTCTATCTCTTCACTCCCTCTTTGAGG 6869  
Db 1807 AGAGGAGACACAGTCCCTGTTGCGTGTAGGATATCCCTAACTTCGGAGCGGAGAC 1866  
Qy 6870 TTTTGGTTA CACCGACCATGTGATAGACACAGCTCAGGGGGAGGTGCGATCTGAG 6929  
Db 1867 GGCAGGACGCTCCACAAAAATGGAGAGGACCACTCTGTGCAGTCCCGTCTGGGGTT 1926  
Qy 6930 GGGGTAACAATTCGTGGGCGTGGGGTACTCCAGTTTATGATCTCCGCGAGAGCGGTA 6989  
Db 1927 CTGATTCAGGGGCGCGTGTGGGGTATTGGAGATCAAAACTCTGGGAGTCCCTTAA 1986  
Qy 6990 CGAGTTCTTCTGTCMAATCTTACGCGCCGATCCACAGGGGAGTACTGAGTTAATAA 7049  
Db 1987 TGASCTCTCTCTCAAAACCTATGCGAGCAGCTCCACTAGTGGCAGCATGCCGTTAATAA 2046  
Qy 7050 GTCCCCGAATTCATCTGTTCTGTAATTTTGTGGAGCTCTTTAGACATTAACCATGTGC 7109  
Db 2047 CACCCCTTATCTTGTCTGCTGCAAGTTTGTACAACTGCTCGAGGAATAAGCCAAATCG 2106  
Qy 7110 TCCGACACTCGGTACGATGAACCGCTCTCTTGTCTGTGTGATGATGAGCACCTGGAGTA 7169  
Db 2107 CCCTAGCGCGGGAACAGGTAACGCTCGCTTGTGCTGTGCTGACCAATATCTGAATGG 2166  
Qy 7170 TCTTAGACAGTATAGGGCCGCTTCACTCTTAAGGTTGCGCTTGTAGAGCGGATGTC 7229  
Db 2167 TCTTTGAAGTATAGGGTCTTCTCAACGTTTAGAGCGGGTACGTGGCAGTCTCGATTGA 2226  
Qy 7230 GGGCATTGACGCTCAGTGTGTTCAATTTACCGACTATATGCGAGTGGCGGAGTAGTGA 7289  
Db 2227 GGGTGGCGACGACAGGGTATCTAACTCTGAGTATCTGATCCAGGACGGGTAAATGAT 2286  
Qy 7290 ACTTGTACATATGATTAAAGATGCTTTTCCGTAATGGCGACTCATGGTCTGTTTTT 7349  
Db 2287 ACCTAAACAGATGTTGTAACAGGTGATCTTTAAGGGGCTTCTCGATGTCTATTGTAATA 2346  
Qy 7350 A-----ACGTTACAGGGTGTAAATAACTCAGAGGCTCTGCGTGGGCAACTT 7398  
Db 2347 CTATGACACGCCACTCTCTCTTAGGGTAAAGAGCTTCGGCGTCTCTGTTGGAAGCTT 2406  
Qy 7399 TTGGGACTCTCCGAAACCGCTCAGGGATAAGCTGTACCAACCGGTGCTTTTATAAAGCA 7458  
Db 2407 CGTGGGCTCTCGACGNACTGAAGGCCAACTCTACAGTGTGTGCTCTTATAAATGA 2466  
Qy 7459 AGCATACGTCAAAAAACATGACCCAGTACCTAAATAAGCTCACAGAGAGTCTGTGCG 7518  
Db 2467 CGCATACAAAACAATCTACGATCCAGTACCTAAATAGAGTGGTGGAGATGTGTGAT 2526  
Qy 7519 ATGTATTGTAATGTAATGTGCTCCCTGGGTCCCATCTCCGGGCTCATCGCGGATCTGAA 7578  
Db 2527 TCGATTAATGAAGAAAGTTCCAGCTGTGGCGCTATGTGCTCTGTTCAGACATTA 2586  
Qy 7579 TCTCTTTAACTGTTCTCCCTCTACCGGGTTCGGGGTTAAACCCCGGGGCGGCCAC 7638  
Db 2587 CATGTTTAACTTTTCTGCTTTTTTCTGCTCTCTGAAATGAGGACCAAGGCGCGGCCG 2646  
Qy 7639 CTGTAACCTCCCGTGGCGCAGAAATGCGCCAGGGATCGTGAAGATTTCTCACCGAGAGGGC 7698  
Db 2647 GTACAACGTGCAATGCGCAGAGGATCCCAAGGCAATTAATCGGATCTCTACGAGAGGAT 2706  
Qy 7699 CTTTGTGTTTACCGAAAAAGATTTTATAGCGTCTGCTGCGAGGGGGTCTGTATACGCC 7758  
Db 2707 CTTATTCTGCACAGAAAAAGGATTTCTGACAGCGCATGCAAGCGGGTGAAGCTCCCTCC 2766



```

QY 7759 CAGTTGGCGCAGATTCTCCAGCAGCTTACCGGAAATGAAGCCCAAGTGTGTTGGGGC 7818
Db 2767 AGCCATATGTAAGTACTACACGAAATATACACTGAAATGAAGCCCAATGCTTGGGGC 2826
QY 7819 GTGGCGTCTCTCATATGTTGAGCGGCGCTATCATGCTATTGCCAGCTCGGTCCCTCGT 7878
Db 2827 CTGGAGCGCAGCTGCTGCAATCGAGGCGCCATATGATATTAACCTCTTCCCTACTGAA 2886
QY 7879 CACTTATATACCTGAGTCCGAGGAGAACTGGAACCTTAGACTCAAGACACTGTGCAA 7939
Db 2887 GCTCTACACACAGTACGATACCGCGCGGTCTCTCTGAGCAGTCCAGGCGCTCTGCT 2946
QY 7939 ACTTGTGTTTCAACCCATCTTTCTTCAAGAAATCTTAGCCCTCTAGCAGTACTGGCAA 7998
Db 2947 TTTGGTTTTTCAACCGTCTACTCTTCCGAGGATTAATGCGCGCTGGAGATCATGACCA 3006
QY 7999 CGGAGATGTTCTGACAAATATTTTACCATCACCGGTACGCGCGAGAGAGGCGCC 8058
Db 3007 GGTACGCTCGCGCTGAAATCTTTTACAGCATCACCGGTCTCTCTGAGAAACCGCGCC 3066
QY 8059 TGTCTGACTGGAAGTACTAGCGGATGACGCTGTCCGGGAAAGCAGCTTGTCCCGATTC 8118
Db 3067 AATTACCACCGCAAGTCACTGGACTGTCTTATCCAGGAAGCGTCTCATGCCAGATC 3126
QY 8119 CTTAATCTGCGAGTATGCGAACCGGGCTTCTCCGGCACCCCTGTTGACCTCAGTAA 8178
Db 3127 TTTAAATTTTGCCTGAGTATGCGAACCGGGCTTCTCCGGCACCCCTGTTGACCTCAGTAA 3186
QY 8179 TGTCTTAGAAATCAGAAATCATCTCAGCGCGCCCAACCCCTGAGTCAATTTGTCTATC 8238
Db 3187 TGTCTGCGCAAAACCGCGTTATTTCTGAGCGCCCTGCTGAGCAGTTTGTCTATG 3246
QY 8239 AATACGACCGCGCTGCTCTGCTGAGTATGCGAACCGGGCTTCTCCGGCACCCCTGTTG 8298
Db 3247 CAACCCCATCCCAACATGCGCGCACCCCTGAGTATGCGAACCGGGCTTCTCCGGCAC 3303
QY 8299 TCCGCGCCATGTTTATGACACAGTGGAAAGCGGTGTCAAGAGCATTCACACAGGCG 8358
Db 3304 TACAGACCGCGGTTTATGATGAGTGGCGCGGTGTCAAGAGCATTCACACAGGCG 8363
QY 8359 ACAGAC 8364
Db 3364 ATCCAC 3369

```

RESULT 5

US-08-757-669A-19  
 ; Sequence 19, Application US/08757669A  
 ; Patent No. 6183751

GENERAL INFORMATION:

APPLICANT: Chang, Yuan  
 APPLICANT: Bohenzky, Roy A.  
 APPLICANT: Russo, James J.  
 APPLICANT: Edelman, Isidore S.  
 APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/757,669A

FILING DATE:

CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 45185-F  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 278-0400  
 TELEFAX: (212) 391-0525  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35100 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-757-669A-19

Query Match

Best Local Similarity 54.4%; Pred. No. 1.3e-204;  
 Matches 1569; Conservative 0; Mismatches 1280; Indels 37; Gaps 5;

```

QY 5511 GGCAGCGGTCAATGCCAATTTCCAGATTTGTTAAATCTGTTCCGAACATACGAGTCTAAAT 5570
Db 489 GACACTGGCAATGATGACTCCACCGTCTCACGTCGAGGTAAATGTCTATGAAT 548
QY 5571 CAAGTATTAACCAACGAAGTTCAAAACAGAGAGCAACCGCGCTGTTTCCACCCGAGCATC 5630
Db 549 GTGCTAGAGCCAGCAGACGTTCAACACAGATGAATCAAGCTAAGCTCCGCGCGGAATC 608
QY 5631 CCGCTCTACAAACCCGAACTCTCTAGGAACGCATCAAAAGCCTTCAGACAAATTCGGA 5690
Db 609 CACATCCACAAGGGGTATTGCTCCGCTGTCTGTTAGTCTGGAATAGAACTCAGA 668
QY 5691 AACGAAACCCCTGTTCTGAGCACCAATCTGTTATATAAACTGCGGTTAAAGACGAGCGGC 5750
Db 669 AAAAGACACTGACCCACCAAGAGAACCTGCGCTCTTGAAGTTGATGAGCCCGCAGA 728
QY 5751 CACACACTGCTGTTGAGGATTACGTTAGTGGCTTCAGGCTTGTGACAGCCCG 5810
Db 729 AAGATGTTCTCCGTTGGGCAAGAGCTTGGGGGGGAGAGATGGCGCTACAGTGGT 788
QY 5811 TAGTGGCTGGCGCGCTCGCGACCTTTGACAGCGCAGCCAGCTTCAGAGCGAGGCA 5870
Db 789 GATTTCTTACCAACGGTCTACATTTGGTGGGACCCACAGGCGTGTTCAGTATCAGCAT 848
QY 5871 GAAGCGCTCTTTGAGTCGCGCATATCTCGGCATAGGAGATTCGCTCGCGTATCTTT 5930
Db 849 AATCTATCTTTGAGTCTATCCAGATCAAGTCTAGTCTGATGCTGTTGCTGGCATTT 908
QY 5931 TAGTGCATATATATCTGTTGGGTAAACACCTCACTCCCTGTAAAGGGGAATAGGTC 5990
Db 909 TGCCCGCATGTACATTTCTCTGCCACATATTTTAAACATCTGTAATCTGGAAGTAGATT 968
QY 5991 GGTCTTTCTATCTGTTTCTCTGAATCTATAGTTACATCTTTAAGTATATAACAAAGACGC 6050
Db 969 CAGTCTGGTGTGAGCCCCCGGGGAGCCAGCGTATGCTTCAGGACCCAGGAGCGC 1028
QY 6051 TAAAAACCCAGATGTGTGACACCTTTTAAAAATCGAAAGCAGCGAGTGTCTATTC 6110
Db 1029 TAAGAACCCCGGTGTCCGCGCTCCGGAACAGACCTCTGGAATACGCTCGTCTGTGAC 1088
QY 6111 CGCGTTTCCATGTCTTCAAGGTAGTTCAGATGAGATCTCCAGTTAGACTCTGTCTCCCT 6170
Db 1089 GAAACCCGATGTGTACCGAATGCCAATCTGTGCGCTCCAGCTCTCAAAATTTTCATC 1148
QY 6171 GTTAAACCCCTTAAGTGAATGAACAGCCATGACCGTGAAGTCTCTGATGTTCTCT 6230
Db 1149 TCCAAATACCCGGAATTTGGGATACACACCTCCATGTTTCACTCATCTAGCTAGGCTCTC 1208
QY 6231 AAAAATAACCCCGAAACAGAGCTTAAATACCAATGACTGTACCCCTTACCCCAAGCCAGC 6290
Db 1209 CCCACCCACCCCAATAGGACCCAGCTACAGCTTATCTCCACTAAATACCCAGGAGCTA 1268

```

QY 6291 CCCGTAATTT-----AGACGAGGTCAGTAACCAACGCTATCCT 6330  
Db 1269 CCGGGACTCATTAAGCCGCCCCAGAAAACGAGTAGCTGGGTGGCAATGACAGCTCCCT 1328  
QY 6331 TTAATAAACCCATAAGTGGAGTTGTAAAGGTAACAGCT-CGTGTATATTTTCAGAGCCCTG 6389  
Db 1329 TTAATAAAGTCAACCTTACTCCGAAGGGGTAGTCTGTTGTGAGAACTACTGTCCAGCAGC 1388  
QY 6390 TCATAAATGGATACCGACGACAAATCAGGTAATTAATAATTTTATTTTCAAGATTACAGGA 6449  
Db 1389 CACAAAATGGCGCAAGATGACAAAGTAAAGATCGACC--TTTATTTATTAAGTGAACA 1446  
QY 6450 GGGCGTGTATACAGTGTAGTGGAGCATATCTCGCAATATCGGCAATGGGAATGGCTGATGAT 6509  
Db 1447 ATGGGTGTATTAATATGTTAGTGGAGCAGAGTTTCGCCAAGCTCTAAGTCCGAACAGT 1506  
QY 6510 CCACACTTATGTCGTCGATCAGGTGCTCTGTGTACTTTTACTAGTCCGATGACCTAAT 6569  
Db 1507 CGGTGTACAGGCTCTTATTAAGTGTTCGGTGTACTTGACAAAGCCGCGAACCCTAGGT 1566  
QY 6570 TTGGCGTGTATGAGTGTGTAACAACCAAAATCTGCCCTCGCTTTTATAGAGCTCTA 6629  
Db 1567 TGGGTCTGTACAGGTCTGTACAGGCAAAAAGGATCGGCGGTGCTTTTCAGAGAGTTA 1626  
QY 6630 TGCAAGCTTTAATCGTCCAAATAGCTTGTCTTAAATAATGCAACTCTGGTACACTTTA 6689  
Db 1627 GGGACGTGCTGATTAATGTGACAAAGCTTCTGCTGTAATGCAACCCTGTCATCTGAA 1686  
QY 6690 TGACAGTGTGTCAAAATAAACAAGATTTAAACACAGTTAAACTCGTCTCGTGTAGT 6749  
Db 1687 CGACAGCTGTCAAAAATAAACAAGGTTCAAGTGCACGTTAAATCTGTATCTCGAAAGT 1746  
QY 6750 CCTCGTATATAACCTCTCAACAGAGAAATAATTTTAAACCAACCTCCAGGACTGAA 6809  
Db 1747 CCTGTAAATGACAGTTTCTACAAAGAAACTTTTACACGCTGGCCATCCACTGAA 1806  
QY 6810 ACAGATGTPACGACAGTCTGTGAAAGTGTCTATCATCTCTTCACTCCCTCCCTTTGAAAG 6869  
Db 1807 AGGAGGAGCACAGCTCCGCTGTGCGTGTGTAGGATATCCCTTAACTTCGGAGCGGAGAC 1866  
QY 6870 TTTTGTGTACACGACCATGTGATAGTAGGACACAGCTCAGGGGGAGTCCGATCGTAG 6929  
Db 1867 GCGCGAGCTCCCAAAATAGGAGGAGGACCACTCTGTGCAATCCGCGTCTGGGTT 1926  
QY 6930 GGGGTAACAATTCGTGGCGGTGGGGTATCTCCAGTTTATAGATCTCCGCGACGACGCTA 6989  
Db 1927 CTGATTCAGGGGCGCGTGTGGGGTATTGGAGAGTCAAAACTCTGGGCACTCCCTTAA 1986  
QY 6990 CGAGTCTTCGTCAAACTTACGACGCGCCGATCCACAGGCGGAGTCTGAGTTAATAA 7049  
Db 1987 TGAGTCTCTCTCAAAACCTATGACGACGCTGCTAGTGGAGCATGCCGTTAATAA 2046  
QY 7050 GTCCCGAATTTCTATCGTTCGTAAATTTGTGGAGCTCCCTTTAGACAGTAAACCATGTTGC 7109  
Db 2047 CACCCCTTATCTGTCTGTGCAAGTTGTACAACTGCTGAGGGAATAAGCAAAATCG 2106  
QY 7110 TCCGACACTCGGTAACGATGAAACGCTCTCTGTCTGTGTGATGACGACCTGGAGTA 7169  
Db 2107 CCCTAGCCGCGGAACCAAGTACGCTGCTTTGTGGTGTGAGCAATATCTGAATGG 2166  
QY 7170 TCTTAGACAGTATAGGGCCCGTTTCCACTCTAAGGGTGTGCGGCTGTGTAGAGCGGATTGC 7229  
Db 2167 TCTTTGAAGGTAAGGCTCTTCTCAACGTTTAGAGGGGTACGTGGCAGTCTGGATTGA 2226  
QY 7230 GGGCATTTGACGCTCAGTGTGTCTATTTACCGGACTATCATCGGAGTGGCGGATAGTGA 7289  
Db 2227 GGGTGGGACGGAAGGATCTTACTCTCTGAAGTATCTGATCCAGGACGGGTAAATGAT 2286  
QY 7290 ACTTGTACATATTAAGATGTTCTTTCCCGTAAATGGGCACTCATGCTCGTGTGTTT 7349  
Db 2287 ACCTAAACAGATGTTGAACAGGTGATCTTTAAGGGGCGCTTCTCCGATGTCAATGTAATAA 2346

QY 7350 A-----ACGTTCAAGGGTGTAAAAAACTCAGAAAGTCTCGTGGCAAACTT 7398  
Db 2347 CTATGACACGCCACTCTCTCTCTAGGGTAAAGAGCTTCGGCGGTCTGTGTGGAAGCTT 2406  
QY 7399 TTGGGACTCTCCGAACCCCTCAGGGATAAGCTGTACCAAGGGTCTCTTTATAAACA 7458  
Db 2407 CGTGGGCTCTCGGACGAACCTGAAGCCCAACTCTACCAAGTGTGTCTCTTTATAAATGA 2466  
QY 7459 AGCATAGCTCAAAAAACATGACCCACGTCAGCTAAATAGCTACAGAAAGTCTGTGCG 7518  
Db 2467 CGCATAGCAAAACATCTAGTCCAGTGACCTAAATAGAGTGGTGAAGATGTGTGCAT 2526  
QY 7519 ATGTATTTAATGAAGATGTCTCTGGTCCCTCTCGGGTCCCATCTCCGGGCTCATCGCGATCTGAA 7578  
Db 2527 TCGAATTTAAGAAAGATGTTCAAAGCTTGGTGGCTATGTGTCTGTCTTTACAGACATTA 2586  
QY 7579 TCTCTTTAACTCTCTCTCTCTACCGGGTTCCTGGGTAAAAACCGGGGGCGCCAC 7638  
Db 2587 CATGTTTAACTTTCTCTCTTTTCTGTCCTCTGATGAGGACAAAGGCGCGCGCG 2646  
QY 7639 CTGTAAAGTCCCTGTCGCAAGATGCGCCACAGGGCATCTGTAGAAATTTCTACCGAGGGC 7698  
Db 2647 GTACAAAGTCCCTGTCGAGAGGATCCCAAGGCAATATTCGATCTCTCACGAGAGGAT 2706  
QY 7699 CCTTGTGTGACGAAAGATGTTCATAGCTGTCTGTCAGCGGGTCTGTATACCGCC 7758  
Db 2707 CTTATTTGCAAGAAAGGCAATTTCTGACGCGCATGACGCGGGTGTGAGCTGCTCTCC 2766  
QY 7759 CCAGTTGGCCAGAGTTCTGACGAGCTCTACGCGAGTCTACGCGAAATGAAGGCCAAAGTTTGGGGC 7818  
Db 2767 AGCCATATGTAAGCTACTACAGAAATATACACTGAATGAAGGCCAAATGCTGGGGC 2826  
QY 7819 GTGGCTGTCTCATATGTTGACAGGGCTCTATCATGGCTATTGCGGACTGTGCTCTCGT 7878  
Db 2827 CTGAGGGGACTGTCTGCAATCGAGGGCCATTAATGATTAATTAATCTCTTCCCTACTGAA 2886  
QY 7879 CACTTATATACCTCGATGCGCGAGGAAAACTGGAACCTTAGACTCAAGACACTGTGMA 7938  
Db 2887 GCTCTACAAACAGCTAGCATACCGCGGGTGTCTCTGAGCAGTCTCAGGGGCTCTGCT 2946  
QY 7939 ACTTGTTTTCAACCCATCTTTCTCAAGAAATCTTAGCCCTATGCGAGTCTACTGCGCAA 7998  
Db 2947 TTTGTTTTCACACCGTCTACTCTTCGAGGATTAATGGCGGCTGGAGATCATACCAA 3006  
QY 7999 CGGGAAGATGTTCTGTGACAAATTAATTTTACATACCGGTACGGCCGAGAGCGGCC 8058  
Db 3007 GGGTCACTCGCCCTTGAAAACTTTTACAGCATCACCGTCTCTGTGTGAAAAACGCGGCC 3066  
QY 8059 TGTGCTGACTGGAAGTACTAGCGGGATGAGTGTCCGGGAAGCAGCTTGTCCCGATTC 8118  
Db 3067 AATTACACCGGCAAGTCACTGGACTGTCTATCAGGAAGCGTCTCATGCGCAGATC 3126  
QY 8119 CTTAATCTCGCATATGCAACCGGGGCTTCTCCGGCAACCCCTGTGTGACTCAGTAA 8178  
Db 3127 TTTAATTTGCAATCTCGGAGCAGGACTGTTGCGGCTTCCATGTGTAGACTCTCAGCGA 3186  
QY 8179 TGTCTTAGAAAAATCCAGAAATCATCTCAGCGCCCAACCCCTGTGACTCAATTTGTCTATC 8238  
Db 3187 TGTGCTGGCAAAACCCCGCTTATTTCTGAGCCCTCTGCCCTGAGCCAGTTTGTCTTAG 3246  
QY 8239 AAACACGCAACCCAGTCTGCTCAGTCACTGAGCAATTAATAGCCAAACCCAGGGCGTTGT 8298  
Db 3247 CAACCCCATCCCAACATGCCGACACCCCTCAGCATCATCCCTTTAAACC---CATCGG 3303  
QY 8299 TCCCGGCCAATGTTTATGGAACACGTTGGAAGCGGTGTACAGAGCATTTCAACACCGGC 8358  
Db 3304 TACAGACCGCGGTTTATAGTACGTGCGAGCCGCTCACAGAAATATGGTGTACACAC 3363  
QY 8359 ACAGAC 8364  
Db 3364 ATCCAC 3369

RESULT 6  
US-09-230-371A-19  
Sequence 19, Application US/09230371A  
Patent No. 6348586  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A  
APPLICANT: Russo, James J  
APPLICANT: Edelman, Isidore S  
APPLICANT: Moore, Patrick S  
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
TITLE OF INVENTION: US/09-230, 371A  
FILE REFERENCE: 45185-G-PCT-US  
CURRENT FILING DATE: 1999-11-17  
CURRENT APPLICATION NUMBER: PCT/US97/13346  
PRIOR FILING DATE: 1997-07-22  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 19  
LENGTH: 35100  
TYPE: DNA  
ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-230-371A-19

Query Match 7.9%; Score 714; DB 4; Length 35100;  
Best Local Similarity 54.4%; Pred. No. 1.3e-204;  
Matches 1569; Conservative 0; Mismatches 1280; Indels 37; Gaps 5;

QY	5511	GGCAGCGCTATTGCCATTTCCAGATTGTTAAATCTGTCGACATACGAGTCTAAAT	5570
DB	489	GACACTGGCCATGATGACTCCAAACCGCTCTCAGTCGAGTAATGTCTCTATGAGAT	548
QY	5571	CAAGTATAACACACACAGTAAGTTCAACAGAGGAGCAACGGCGCTGGTTCCACCGCGACATC	5630
DB	549	GTGGTAGAGCCAGCAGAGCTTCAACACGATGAATCAAGCTAAGCTCCCGCGGAATC	608
QY	5631	CCGCTCACAAACCCGAACTCTTAGAAACGACATCAAGAGCTTCAGACAAATTCGA	5690
DB	609	CACATCCACAAAGGGATTGCTCCGGTGTCTGTATTAGTCTGGAATAGAAATCTCAG	668
QY	5691	AACCGAAACCTAGTTCTGAGCAACATCTGTTTAAATCAACGCGTTAAAGAGACGAGCGC	5750
DB	669	AAAGACATGACCCACAGAGAGACTCGCGCTCTTGCAAAATGATGAGCCCCGAGA	728
QY	5751	CACACTGCTGTTGTGAGATTTCAGCTTAGTGGCTGCGAGCTGTGACAGCCCG	5810
DB	729	AAGAAATGTCTCCGTTGGACAAAGAGCTTGGGGGGCAGAGATGGCGCTACAGTGGGT	788
QY	5811	TAGTGTGCTGGCGGCTCGGACCTTTGACAGCCGACCGCAGTTCCAGAGCGAGGCA	5870
DB	789	GAATTTCTTACACCGGTCTACATTTGGTGGCACCCACAGGCTGTTTCCAGTATCAGAT	848
QY	5871	GAAGCGCTCTTTGAGTGGGCCATATCTCGGCCATAGGAGATTGCGTCCGCTGATCTTT	5930
DB	849	AAATCTATCTTTGAGTCAATCCAGATCAAAATCATGTGATGATGCTGTTCCCTGGCATTT	908
QY	5931	TAGTGTGATATATATCTGTTGGGTAAACAACTCACTCCCTGTAAAGGGGATAGGTC	5990
DB	909	TGCCCGCATGATACATTTCTGTCACATATTTTAAACATCTGTAATCTGGAAGTGAAT	968
QY	5991	CGTTCTTCTATCTGTTTCTCTGAATCTATAGTTACATCTTTAAGTATAACCAAGACGC	6050
DB	969	CAGTCTGTTGAGCCCCCGGGGAAGCCAGCGTATGCTTCAGGACCCAGGAGCGC	1028
QY	6051	TAAAAACCCAGGATGGTGCACCTTTAAATAATCGCAAGACCCGATGGCTATC	6110
DB	1029	TAAAGAACCCGGGTGTCGCGCTCCGGAAACAGACCTCTGAGAAATACGCTCGGTCTTGAC	1088
QY	6111	CGCGTTTCCATGTCGTTCAAAAGGTAGTCACGATGGATCTCCAGTTAGACTCTGTCTCCCT	6170
DB	1089	GAACCCGATGTTGATCCGAATGCCAATCTGTGGCCCTCCAGCTCTCACAAATTTTCATC	1148

QY	6171	GTTAACACCCCTTAACCTGAATGGAACAGCCATGACACCGTAACCTTCCTGATGTCCT	6230
DB <td>1149<td>TCCATATACCCGGAATTTGGGATACACACTCCATGTTCCAGTCACATGTAAGCTGGTCTC<td>1208</td></td></td>	1149 <td>TCCATATACCCGGAATTTGGGATACACACTCCATGTTCCAGTCACATGTAAGCTGGTCTC<td>1208</td></td>	TCCATATACCCGGAATTTGGGATACACACTCCATGTTCCAGTCACATGTAAGCTGGTCTC <td>1208</td>	1208
QY <td>6231<td>AAACTAACCCCGAAACAGAGCTAAATACCAATGACTGTACCCCTTACCCCAAGCCACGC<td>6290</td></td></td>	6231 <td>AAACTAACCCCGAAACAGAGCTAAATACCAATGACTGTACCCCTTACCCCAAGCCACGC<td>6290</td></td>	AAACTAACCCCGAAACAGAGCTAAATACCAATGACTGTACCCCTTACCCCAAGCCACGC <td>6290</td>	6290
DB <td>1209<td>CCACCCCAACCCCATAGGACCCAGCTATCTCTCCATTAATACAGGACGCTA<td>1268</td></td></td>	1209 <td>CCACCCCAACCCCATAGGACCCAGCTATCTCTCCATTAATACAGGACGCTA<td>1268</td></td>	CCACCCCAACCCCATAGGACCCAGCTATCTCTCCATTAATACAGGACGCTA <td>1268</td>	1268
QY <td>6291<td>CCCGGTACTATT-----AGACCAGGTGAGTAACCAACCGTATCTCT<td>6330</td></td></td>	6291 <td>CCCGGTACTATT-----AGACCAGGTGAGTAACCAACCGTATCTCT<td>6330</td></td>	CCCGGTACTATT-----AGACCAGGTGAGTAACCAACCGTATCTCT <td>6330</td>	6330
DB <td>1269<td>CCGCGACTCATTAAGCCCGCCAGAAACAGAGTGTGGTGGCAATGACACGTCCTCT<td>1328</td></td></td>	1269 <td>CCGCGACTCATTAAGCCCGCCAGAAACAGAGTGTGGTGGCAATGACACGTCCTCT<td>1328</td></td>	CCGCGACTCATTAAGCCCGCCAGAAACAGAGTGTGGTGGCAATGACACGTCCTCT <td>1328</td>	1328
QY <td>6331<td>TTAAAAACCCATACGTGGAGTTTGAAGTTAAACAGCT-CGTGTATATTTTCAGACGCTG<td>6389</td></td></td>	6331 <td>TTAAAAACCCATACGTGGAGTTTGAAGTTAAACAGCT-CGTGTATATTTTCAGACGCTG<td>6389</td></td>	TTAAAAACCCATACGTGGAGTTTGAAGTTAAACAGCT-CGTGTATATTTTCAGACGCTG <td>6389</td>	6389
DB <td>1329<td>TTAAAAAGTCAACCTTACTCCGCAAGGGGTAGTCTGTGTGAGAAATCTGTCCAGGACG<td>1388</td></td></td>	1329 <td>TTAAAAAGTCAACCTTACTCCGCAAGGGGTAGTCTGTGTGAGAAATCTGTCCAGGACG<td>1388</td></td>	TTAAAAAGTCAACCTTACTCCGCAAGGGGTAGTCTGTGTGAGAAATCTGTCCAGGACG <td>1388</td>	1388
QY <td>6390<td>TCATAAATGATACCCAGCAGCATCAGGTAATTAACCTTTTATTTTCAAGATTCAGGA<td>6449</td></td></td>	6390 <td>TCATAAATGATACCCAGCAGCATCAGGTAATTAACCTTTTATTTTCAAGATTCAGGA<td>6449</td></td>	TCATAAATGATACCCAGCAGCATCAGGTAATTAACCTTTTATTTTCAAGATTCAGGA <td>6449</td>	6449
DB <td>1389<td>CACAAAATGGCGAAGATGACAGGTAAGATCGACC--TTTTTATGTTATCTGAACA<td>1446</td></td></td>	1389 <td>CACAAAATGGCGAAGATGACAGGTAAGATCGACC--TTTTTATGTTATCTGAACA<td>1446</td></td>	CACAAAATGGCGAAGATGACAGGTAAGATCGACC--TTTTTATGTTATCTGAACA <td>1446</td>	1446
QY <td>6450<td>GGCGGTGTTTACAGTGTGTAGGTGGAGCATATCTCGCAATGGGAATGGCTGATGAT<td>6509</td></td></td>	6450 <td>GGCGGTGTTTACAGTGTGTAGGTGGAGCATATCTCGCAATGGGAATGGCTGATGAT<td>6509</td></td>	GGCGGTGTTTACAGTGTGTAGGTGGAGCATATCTCGCAATGGGAATGGCTGATGAT <td>6509</td>	6509
DB <td>1447<td>ATGCGTGTTTACAAATGGTGTAGGTGGAGCAGAGTTCGCCAAGCTCTACGTCCGAAAGT<td>1506</td></td></td>	1447 <td>ATGCGTGTTTACAAATGGTGTAGGTGGAGCAGAGTTCGCCAAGCTCTACGTCCGAAAGT<td>1506</td></td>	ATGCGTGTTTACAAATGGTGTAGGTGGAGCAGAGTTCGCCAAGCTCTACGTCCGAAAGT <td>1506</td>	1506
QY <td>6510<td>CCACACTAGTGTCCGATCAGGTGCTCTGTGACTTTTACTAGTCCGATGAACCTAAAT<td>6569</td></td></td>	6510 <td>CCACACTAGTGTCCGATCAGGTGCTCTGTGACTTTTACTAGTCCGATGAACCTAAAT<td>6569</td></td>	CCACACTAGTGTCCGATCAGGTGCTCTGTGACTTTTACTAGTCCGATGAACCTAAAT <td>6569</td>	6569
DB <td>1507<td>CGGTGTGAGGCTCTTATTAAAGTGTTCGGTGTACTTGACCAAGACCCGCGAATCTAGT<td>1566</td></td></td>	1507 <td>CGGTGTGAGGCTCTTATTAAAGTGTTCGGTGTACTTGACCAAGACCCGCGAATCTAGT<td>1566</td></td>	CGGTGTGAGGCTCTTATTAAAGTGTTCGGTGTACTTGACCAAGACCCGCGAATCTAGT <td>1566</td>	1566
QY <td>6570<td>TTGGCGTGTGTAGTGTACCAACCAAAACTCCCTGGCTCCGCTTTTGAAGCTCTA<td>6629</td></td></td>	6570 <td>TTGGCGTGTGTAGTGTACCAACCAAAACTCCCTGGCTCCGCTTTTGAAGCTCTA<td>6629</td></td>	TTGGCGTGTGTAGTGTACCAACCAAAACTCCCTGGCTCCGCTTTTGAAGCTCTA <td>6629</td>	6629
DB <td>1567<td>TGGGTCTGTACAGTGTGTACCAAGGATCGGGCGGTCTTTTCAGGAGAGTTA<td>1626</td></td></td>	1567 <td>TGGGTCTGTACAGTGTGTACCAAGGATCGGGCGGTCTTTTCAGGAGAGTTA<td>1626</td></td>	TGGGTCTGTACAGTGTGTACCAAGGATCGGGCGGTCTTTTCAGGAGAGTTA <td>1626</td>	1626
QY <td>6630<td>TGCAAGCTTTAATCGTGTCCATAGCTTGTCTTTAAATGCAACTCTGTGATCTTTA<td>6699</td></td></td>	6630 <td>TGCAAGCTTTAATCGTGTCCATAGCTTGTCTTTAAATGCAACTCTGTGATCTTTA<td>6699</td></td>	TGCAAGCTTTAATCGTGTCCATAGCTTGTCTTTAAATGCAACTCTGTGATCTTTA <td>6699</td>	6699
DB <td>1627<td>GGGAGCTGTGATTTATGTGGCAAGCTTCTGCTGTAATGACCGCTGTGTACTCTGAA<td>1686</td></td></td>	1627 <td>GGGAGCTGTGATTTATGTGGCAAGCTTCTGCTGTAATGACCGCTGTGTACTCTGAA<td>1686</td></td>	GGGAGCTGTGATTTATGTGGCAAGCTTCTGCTGTAATGACCGCTGTGTACTCTGAA <td>1686</td>	1686
QY <td>6690<td>TGACAGTGTCTCAAAAAAACAAGATTTAAAAACAGTTTAAACTGCGTGTCTGGTGT<td>6749</td></td></td>	6690 <td>TGACAGTGTCTCAAAAAAACAAGATTTAAAAACAGTTTAAACTGCGTGTCTGGTGT<td>6749</td></td>	TGACAGTGTCTCAAAAAAACAAGATTTAAAAACAGTTTAAACTGCGTGTCTGGTGT <td>6749</td>	6749
DB <td>1687<td>CGACAGCTGTCCAAAAAACAAGTTTCACTGACGTTAAATCTGTATCTCTGAAAGT<td>1746</td></td></td>	1687 <td>CGACAGCTGTCCAAAAAACAAGTTTCACTGACGTTAAATCTGTATCTCTGAAAGT<td>1746</td></td>	CGACAGCTGTCCAAAAAACAAGTTTCACTGACGTTAAATCTGTATCTCTGAAAGT <td>1746</td>	1746
QY <td>6750<td>CCTGTATATAAATCTCTCAACAAAGAAATTTTAAACCAAACTGCCAGTACTGAA<td>6809</td></td></td>	6750 <td>CCTGTATATAAATCTCTCAACAAAGAAATTTTAAACCAAACTGCCAGTACTGAA<td>6809</td></td>	CCTGTATATAAATCTCTCAACAAAGAAATTTTAAACCAAACTGCCAGTACTGAA <td>6809</td>	6809
DB <td>1747<td>CCTCGTAATGACAGTTTCTACCAAGAAAACTTTTACCAACGCTGCCATCCTGAA<td>1806</td></td></td>	1747 <td>CCTCGTAATGACAGTTTCTACCAAGAAAACTTTTACCAACGCTGCCATCCTGAA<td>1806</td></td>	CCTCGTAATGACAGTTTCTACCAAGAAAACTTTTACCAACGCTGCCATCCTGAA <td>1806</td>	1806
QY <td>6810<td>ACGATGTGTACGACAGGTCTGTGAAAGTGTCTATCATCTCTTCACTCCCTTTGAAG<td>6869</td></td></td>	6810 <td>ACGATGTGTACGACAGGTCTGTGAAAGTGTCTATCATCTCTTCACTCCCTTTGAAG<td>6869</td></td>	ACGATGTGTACGACAGGTCTGTGAAAGTGTCTATCATCTCTTCACTCCCTTTGAAG <td>6869</td>	6869
DB <td>1807<td>AGGAGGAGCACAGTCCCGTGTGCGTTGTAGGATATCCCTAACCTTCGAGCGGAGAC<td>1866</td></td></td>	1807 <td>AGGAGGAGCACAGTCCCGTGTGCGTTGTAGGATATCCCTAACCTTCGAGCGGAGAC<td>1866</td></td>	AGGAGGAGCACAGTCCCGTGTGCGTTGTAGGATATCCCTAACCTTCGAGCGGAGAC <td>1866</td>	1866
QY <td>6870<td>TTTTGGTTTACCCGACCATGTGATAGGCAACAGCTCAGGGGAGGTGCGATCGTGAG<td>6929</td></td></td>	6870 <td>TTTTGGTTTACCCGACCATGTGATAGGCAACAGCTCAGGGGAGGTGCGATCGTGAG<td>6929</td></td>	TTTTGGTTTACCCGACCATGTGATAGGCAACAGCTCAGGGGAGGTGCGATCGTGAG <td>6929</td>	6929
DB <td>1867<td>GGCGGACGCTCCACAAAAATGGAGAGGCAACCACTCTGTGCACTCCGCTGTGGGTT<td>1926</td></td></td>	1867 <td>GGCGGACGCTCCACAAAAATGGAGAGGCAACCACTCTGTGCACTCCGCTGTGGGTT<td>1926</td></td>	GGCGGACGCTCCACAAAAATGGAGAGGCAACCACTCTGTGCACTCCGCTGTGGGTT <td>1926</td>	1926
QY <td>6930<td>GGGTAAATTCGTGGCGTGGGTACTCCAGTTTGTAGTCTCCGCGCAGGACCGTA<td>6989</td></td></td>	6930 <td>GGGTAAATTCGTGGCGTGGGTACTCCAGTTTGTAGTCTCCGCGCAGGACCGTA<td>6989</td></td>	GGGTAAATTCGTGGCGTGGGTACTCCAGTTTGTAGTCTCCGCGCAGGACCGTA <td>6989</td>	6989
DB <td>1927<td>CTGATTCAGGGGCGCGTGTGGGGTATTGGAGTCAAACTCTGGGCTCCCTTAA<td>1986</td></td></td>	1927 <td>CTGATTCAGGGGCGCGTGTGGGGTATTGGAGTCAAACTCTGGGCTCCCTTAA<td>1986</td></td>	CTGATTCAGGGGCGCGTGTGGGGTATTGGAGTCAAACTCTGGGCTCCCTTAA <td>1986</td>	1986
QY <td>6990<td>CGAGTTCTTCGTCAAATCTTACGAGCCGATCCACAGGGGAGTACTGAGTTAAATA<td>7049</td></td></td>	6990 <td>CGAGTTCTTCGTCAAATCTTACGAGCCGATCCACAGGGGAGTACTGAGTTAAATA<td>7049</td></td>	CGAGTTCTTCGTCAAATCTTACGAGCCGATCCACAGGGGAGTACTGAGTTAAATA <td>7049</td>	7049
DB <td>1987<td>TGAGCTCTCTCAAACCTATGACAGCGCTCCACTAGTGGCAGCATGCCCTTAAATA<td>2046</td></td></td>	1987 <td>TGAGCTCTCTCAAACCTATGACAGCGCTCCACTAGTGGCAGCATGCCCTTAAATA<td>2046</td></td>	TGAGCTCTCTCAAACCTATGACAGCGCTCCACTAGTGGCAGCATGCCCTTAAATA <td>2046</td>	2046
QY <td>7050<td>GTCCCGGATCTATCTGTTTGTGAGTCTCTTTAGACAGTAAACATGTTGC<td>7109</td></td></td>	7050 <td>GTCCCGGATCTATCTGTTTGTGAGTCTCTTTAGACAGTAAACATGTTGC<td>7109</td></td>	GTCCCGGATCTATCTGTTTGTGAGTCTCTTTAGACAGTAAACATGTTGC <td>7109</td>	7109
DB <td>2047<td>CACCCCTTATCTGTGTGTTGCCAGTTTGTACAATCTGTGAGGGAATAGCCAAATTCG<td>2106</td></td></td>	2047 <td>CACCCCTTATCTGTGTGTTGCCAGTTTGTACAATCTGTGAGGGAATAGCCAAATTCG<td>2106</td></td>	CACCCCTTATCTGTGTGTTGCCAGTTTGTACAATCTGTGAGGGAATAGCCAAATTCG <td>2106</td>	2106
QY <td>7110<td>TCGACACTGCGGTACGATGAACCCCTCTCTCTGTGTGTGTGATGAGCACTGGAGTA<td>7169</td></td></td>	7110 <td>TCGACACTGCGGTACGATGAACCCCTCTCTCTGTGTGTGTGATGAGCACTGGAGTA<td>7169</td></td>	TCGACACTGCGGTACGATGAACCCCTCTCTCTGTGTGTGTGATGAGCACTGGAGTA <td>7169</td>	7169
DB <td>2107<td>CCCTAGCCCGGGAACAGGTACGGCTCGCTTGTGCGTGTGGACCAATATCTGAATGG<td>2166</td></td></td>	2107 <td>CCCTAGCCCGGGAACAGGTACGGCTCGCTTGTGCGTGTGGACCAATATCTGAATGG<td>2166</td></td>	CCCTAGCCCGGGAACAGGTACGGCTCGCTTGTGCGTGTGGACCAATATCTGAATGG <td>2166</td>	2166
QY <td>7170<td>TCTTAGACAGGTATAGGCGGCTTCTCACTCTAAGGGTTGCCGTTGTGAGCGGATTCG<td>7229</td></td></td>	7170 <td>TCTTAGACAGGTATAGGCGGCTTCTCACTCTAAGGGTTGCCGTTGTGAGCGGATTCG<td>7229</td></td>	TCTTAGACAGGTATAGGCGGCTTCTCACTCTAAGGGTTGCCGTTGTGAGCGGATTCG <td>7229</td>	7229
DB <td>2167<td>TCTTTGCAAGTATAGGGTCTTCTCAAGTTTAGAGCGGGTACGTGGCAGTCTGATGTA<td>2226</td></td></td>	2167 <td>TCTTTGCAAGTATAGGGTCTTCTCAAGTTTAGAGCGGGTACGTGGCAGTCTGATGTA<td>2226</td></td>	TCTTTGCAAGTATAGGGTCTTCTCAAGTTTAGAGCGGGTACGTGGCAGTCTGATGTA <td>2226</td>	2226
QY <td>7230<td>GGGCATTGACGCTCAGTGTGTGTTTCACTCCGACTATCATCCGAGTGGCGGATAGTGA<td>7289</td></td></td>	7230 <td>GGGCATTGACGCTCAGTGTGTGTTTCACTCCGACTATCATCCGAGTGGCGGATAGTGA<td>7289</td></td>	GGGCATTGACGCTCAGTGTGTGTTTCACTCCGACTATCATCCGAGTGGCGGATAGTGA <td>7289</td>	7289



Db 97132 GTATATGAGTAAGTTACACGGCGCAATTTAAACTA 97165

## RESULT 8

US-09-661-596A-76  
; Sequence 76, Application US/09661596A  
; Patent No. 6528066  
; GENERAL INFORMATION:  
; APPLICANT: Grose, Charles  
; APPLICANT: Santos, Richard  
; TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE  
; FILE REFERENCE: 140,001,010  
; CURRENT APPLICATION NUMBER: US/09/661,596A  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 76  
; LENGTH: 124884  
; TYPE: DNA  
; ORGANISM: Varicella zoster  
US-09-661-596A-76

Query Match 1.7%; Score 152.4; DB 4; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1e-33;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 585 CAATATATTGTAATCGACGAGCGCGGAAACCCCTCTCTTACATATCTACACCGCGT 644  
Db 96706 CAACGTTATGTTATGATGAGCAGGATTCCTAGGCGTCATATCTCAGCGCGTGT 96765  
QY 645 GTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCGCGGCGT 704  
Db 96766 TTACTGTTGGTGGCTTTTGAATGCTATATATCAAGCCCTCAGTACATATAACCGTGA 96825  
QY 705 TCGGTGCATAGTATCGTTCGCTTCCGCAACGACAGAAACCGCTTCAGTCAACTTACA 764  
Db 96826 ACCGTGTAGTATCGTTCGCTTCCGCAACGACAGAAACCGTTCAGTCAACTTACA 96885  
QY 765 CCACGGGACGCAAAAGACGAGATATCGTTCGCGAGAAACATCCTAACATTCATGATAG 824  
Db 96886 ACATGACATGACGGCTTCACACGTAACCTCTAGTGAATAATATCTCAGCTATATACTG 96945  
QY 825 GAAGAGTCTGATCTGAGTACACTTGGAGAGAACTGGCGCTGTTTATAACAA 884  
Db 96946 CAATCAAACTCTGGCTCAATATCTAACATCTCACAATCTGGGCAATCTTTTATAA 97005  
QY 885 CAAGCGCTGCACCGATCTGCAGTTCGGACACCTTCTTAAATTTTAGAGTATATCTTCC 944  
Db 97006 CAAACGATGTCAGAGGACGATTTTGAATCTTTTAAACGCTTGGTACGGCTACC 97065  
QY 945 CATTCTCGACGAGTCATGAGTTACGTAGACAGATTTGCTTCTTAAAGTAAGATTAT 1004  
Db 97066 TATTACCGAGCAGCATGGCGCTCTGGTGCATACATTTGTTACCTGCATCCTATATA 97125  
QY 1005 GGATCCTTTAGATATACATTTGGCTGGACCGACTCTTTTGTCAATAGCGGTAAGGC 1064  
Db 97126 CAATCCTGTAATCTTCCGCGATGGACGCGTCTGTTTGTGATAGGAGTGAGCGC 97185  
QY 1065 GTATCTACAAACCTACACATGCTTAAAGCTA 1098  
Db 97186 GTATATGATAGTTACACGGCGCAATTTAAACTA 97219

## RESULT 9

US-09-913-514-1  
; Sequence 1, Application US/09913514  
; Patent No. 6653069  
; GENERAL INFORMATION:  
; APPLICANT: GOMI, Yasuyuki  
; APPLICANT: SUNAMACHI, Hiroki

; APPLICANT: TAKAHASHI, Michiaki  
; APPLICANT: YAMAMOTO, Koichi  
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vacci  
; FILE REFERENCE: 0216-0454P  
; CURRENT APPLICATION NUMBER: US/09/913,514  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: PCT/JP01/00678  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: JP 2000-62734  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 124884  
; TYPE: DNA  
; ORGANISM: Varicella virus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(124884)  
; OTHER INFORMATION: Dumas Strain  
US-09-913-514-1

Query Match 1.7%; Score 152.4; DB 4; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1e-33;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 585 CAATATATTGTAATCGACGAGCGCGGAAACCCCTCTCTTACATATCTACACCGCGT 644  
Db 96706 CAACGTTATGTTATGATGAGCAGGATTCCTAGGCGTCATATCTCAGCGCGTGT 96765  
QY 645 GTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCGCGGCGT 704  
Db 96766 TTACTGTTGGTGGCTTTTGAATGCTATATATCAAGCCCTCAGTACATATAACCGTGA 96825  
QY 705 TCGGTGCATAGTATCGTTCGCTTCCGCAACGACAGAAACCGCTTCAGTCAACTTACA 764  
Db 96826 ACCGTGTAGTATCGTTCGCTTCCGCAACGACAGAAACCGTTCAGTCAACTTACA 96885  
QY 765 CCACGGGACGCAAAAGACGAGATATCGTTCGCGAGAAACATCCTAACATTCATGATAG 824  
Db 96886 ACATGACATGACGGCTTCACACGTAACCTCTAGTGAATAATATCTCAGCTATATACTG 96945  
QY 825 GAAGAGTCTGATCTGAGTACACTTGGAGAGAACTGGCGCTGTTTATAACAA 884  
Db 96946 CAATCAAACTCTGGCTCAATATCTAACATCTCACAATCTGGGCAATCTTTTATAA 97005  
QY 885 CAAGCGCTGCACCGATCTGCAGTTCGGACACCTTCTTAAATTTTAGAGTATATCTTCC 944  
Db 97006 CAAACGATGTCAGAGGACGATTTTGAATCTTTTAAACGCTTGGTACGGCTACC 97065  
QY 945 CATTCTCGACGAGTCATGAGTTACGTAGACAGATTTGCTTCTTAAAGTAAGATTAT 1004  
Db 97066 TATTACCGAGCAGCATGGCGCTCTGGTGCATACATTTGTTACCTGCATCCTATATA 97125  
QY 1005 GGATCCTTTAGATATACATTTGGCTGGACCGACTCTTTTGTCAATAGCGGTAAGGC 1064  
Db 97126 CAATCCTGTAATCTTCCGCGATGGACGCGTCTGTTTGTGATAGGAGTGAGCGC 97185  
QY 1065 GTATCTACAAACCTACACATGCTTAAAGCTA 1098  
Db 97186 GTATATGATAGTTACACGGCGCAATTTAAACTA 97219

## RESULT 10

US-09-651-656-36/c  
; Sequence 36, Application US/09651656  
; Patent No. 6340566  
; GENERAL INFORMATION:  
; APPLICANT: MCCUTHEN-MALONEY, SANDRA  
; APPLICANT: LIVERMORE NATIONAL LABORATORY  
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE  
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,  
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES

Query Match	1.5%	Score	137.6;	DB 4;	Length	1532;			
Best local Similarity	57.5%;	Pred. No.	1.2e-30;						
Matches	267;	Conservative	0;	Mismatches	194;	Indels	3;	Gaps	1;
Qy	3638	AGGGGAGACGGATGGCGCGCTTTGAGACGAGATGCTCTGTCGGTTTATCAGCGTCGCC	3697						
Db	1104	AGCGGCGACGGATGGCGTGCTTCAGTACATGATGCGGTGCTTATCTATAATCGCCCT	1045						
Qy	3698	TTGCTATAGCCTTTCTTCCCAACAGATAAACGCACTCTGAAGCTTGGCAGAAAGG	3757						
Db	1044	TTCTTTTGGCATGCGATCCCCAACAAACACACGCTTCCTGGGATGCTGTTAAATC	985						
Qy	3758	CAGCTTATTATGATGCTGGTGAACCAATCCAGCCAAAGTGTGGCGTGTACCCCGCTTC	3817						
Db	984	AGGCTGATCACTTATCGGTGAGGTTTCCAGCCGAGGCTGGCGTGGGATGGGCTGA	925						
Qy	3818	CTCTTTTCCACCGTCAAGATGGTGTATTAAGCAGAACTCCCGTTTGGCCCAACAGTCT	3877						
Db	924	CTTGGCGGTACCGTCAACACAGTATTAGTAGCAGAAACGCGCTGACGCGGCCAGCTTTCA	865						
Qy	3878	AAGCACCGGTGAGAAAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAAAATA	3937						
Db	864	AGATACCATGATTAGGCGGGTGAAGCCCGGNATAGTATTTTCCAGCTCTTTATACATTA	805						
Qy	3938	TTTTTGAGGCTTGGAGGTATCTGTAGTCCGGAGCAACGCTAAAGGCTAGTCCGGTGGCT	3997						
Db	804	TTCAATAATGACGGAGGAATGGCAATTCGCGGACGAAACGGAATAATGCCAGCATCGCC	745						
Qy	3998	TGACCC--GCGATGTAAGGATCCTCGGCAAGGATCAGCATTTTATATCTCTCCGGTTCG	4054						
Db	744	TGTCGCGGTCCGTTGATAGGATCCTCGGCGAGANTCACCACCTTTAACGTACCCAACTCT	685						
Qy	4055	CAACAATACGACCAACACATAATTCCTATCTATCGGTGGATAAAT	4098						
Db	684	GTAAAGCGGAACCGGTAAAGACATCTTTTGTGGTGGGTAGAT	641						

RESULT 12  
US-09-489-039A-4593/c  
; Sequence 4593, Application US/09489039A  
; Patent No. 5610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Braton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBS  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342

```

; PRIOR FILLING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4593
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4593

Query Match      1.5%;      Score 134;      DB 4;      Length 723;
Best Local Similarity 56.8%;      Pred. No. 8.2e-10;
Matches 267;      Conservative 0;      Mismatches 200;      Indels 3;      Gaps 1;

Qy 3632 GTAGCCAGGGGACAGCGATGGCGCGCTTGAGAGCAGATGCTCTGTCCGTTTATCAGC 3691
Db 611 GCAGAGCGGTGATGGATTCGGCGCCCTTCAGCACGCAGTGGCGTCAGATCGATATC 552

Qy 3692 GTGCGCTTGTCTATAGCTTTCTTCCCAACAGATATAAACCGACCTCTGAAGCTTGGCA 3751
Db 551 GCGCCTTTCTCTGGCGATGGGAGGCCCAACAGAGAAAACCTACTCCCTCGCAATGCTCA 492

Qy 3752 GAAAGCAGCTTATTATGTAGCTGGTGAACCAATCCAGCCAGGTGGCGTGTGACCCC 3811
Db 491 TTGATTAAAGCGATACCTTATTCGTTAAAGGTCTCCAGCCAGACTGGCGTGCAGTGCC 432

```



3812 GCCTTCCTCTTTCCACCGTCAAGATGGTGTAAAGCAGAACTCCCGCTTTGGCCAA 3871  
Db |||||  
431 GCCTGTCGGCGCGGACGGTCAAGACCGTAATTCACGACGACGACCGCTGTCGCGCCAG 372  
Qy |||||  
3872 CAGTCTAAGCAACCGGTGAGAGGACGGTGAACCCAGGATCAGTATGCGCTATCTTTA 3931  
Db |||||  
371 CTTTCGAGATAACCGTGGGTGGGCGGGTAAAGCGCGGGATCGTGGCTCCAGCTCTTTA 312  
Qy |||||  
3932 AAAATATTTTTCAGGCTTCGAGGTATCTGTAGTCCGAGCAACGCTAAAGCTAGTCCG 3991  
Db |||||  
311 TACATATTCAGCAACGAGCGGGATCGGACGCGCCGCGAATCTGAGACGCGCAGACCG 252  
Qy |||||  
3992 GTGGCTTACC---GGATGTAAGATCTGCGCCAGGATCAGACTTTAAATCTCC 4048  
Db |||||  
251 TGGGCTGACCGCGCGCGGTGATAGGGGCTGCGCGAGATGACGACTTGCATCGCCC 152  
Qy |||||  
4049 GGTTCGCAATATACGACCAACACATATCTATCTATCGCTGGATAAT 4098  
Db |||||  
191 AGTTCAGTAAGCGGAACGCAATGAGACGCTTTTTCGGTGTAGAT 142

## RESULT 13

US-09-289-254-3  
; Sequence 3, Application US/09289254  
; Patent No. 6299882  
; GENERAL INFORMATION:  
; APPLICANT: Junker, David E.  
; TITLE OF INVENTION: No. 6299882el Recombinant and Mutant Herpesviruses  
; FILE REFERENCE: SY0994  
; CURRENT APPLICATION NUMBER: US/09/289,254  
; CURRENT FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 7316  
; TYPE: DNA  
; ORGANISM: Marek's Disease Virus 1  
US-09-289-254-3

Query Match 1.4%; Score 125.8; DB 3; Length 7316;  
Best Local Similarity 53.3%; Pred. No. 1.4e-26;  
Matches 265; Conservative 0; Mismatches 232; Indels 0; Gaps 0;  
Qy 586 AATATATTGTATCGAGAGCGCGACCTGCTCTTACATACACACCGCGGTG 645  
Db |||||  
6435 AATATAATCATCATGATGAGCGCGATGTTGGGAAATATTACTACGCTATTGTA 6494  
Qy |||||  
646 TTCTTTTACTGTTCTCTAAACAGCTGGCTAAATACCCCTCTTACCGCAGGGCGGTT 705  
Db |||||  
6495 TACTGTTGGCTTACCAATGCCGTGTATCGTACTCTCCAAATATAAAGAGGGTAAAA 6554  
Qy |||||  
706 CCGTGCAATAGTATCGTGGTTGCGCAACGAGCAACGCGTCCAGTCAACTACAAC 765  
Db |||||  
6555 CCAGTATTAAATATGTGCGTGGCTACACAAAGATTCTTGGATGCAATTGAA 6614  
Qy |||||  
766 CACGGAGCCAAAAGACCGAGATATCGTGGCGAGAACATCCPAACATTGATAGGG 825  
Db |||||  
6615 CATAGCAAACTGAGGTGAATGTAAGAATAAGCGAAATATCTGACTTATATATATGC 6674  
Qy |||||  
826 AAGAGGTGCTATCTGAGTACGTACACTTGGAGGAACTGGCGCTGTTTATAAACAAC 895  
Db |||||  
6675 AATCAAACTTACGATCATATTAGACTTATCCAAATATTGGCAATATTATTAATAT 6734  
Qy |||||  
886 AAGCGCTGACCGATCTGCGATTCGGAACCTTCTAAAAATTTTAGAGTATAATCTTCCC 945  
Db |||||  
6735 AAACGCTGTACAGAAACCGAATTTGGAGATCTTCTCAAAACATTTGGAATATGACTTCTC 6794  
Qy |||||  
946 ATTCTGAGCAATCATGATGTTACGTACGACATTTGCTTCTCTAAAGTAAAGTATATG 1005  
Db |||||  
6795 ATAACGAGAGATGCGCGAATGCGAATACCTTTGTTGTCGGAAGCATTTATCAT 6854  
Qy |||||  
1006 GATCCTTTAGAAATACATGCTGGACCGGACTCTTTTGTGCATACGAGGTAAGGGCG 1065  
Db |||||

Db 6855 AATCCGGGAGACCTACCTGGCTGGACAGCTTTATATTCATCTCATTAAGGAAGTACAGTACA 6914  
Qy |||||  
1066 TATCTAAACAACTTACA 1082  
Db |||||  
6915 TATATGATGATGATTACA 6931  
RESULT 14  
US-09-557-884-1/c  
; Sequence 1, Application US/09557884  
; Patent No. 6506581  
; GENERAL INFORMATION:  
; APPLICANT: Fielesmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
; COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995  
; ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1  
Query Match 1.3%; Score 117.6; DB 4; Length 1830121;  
Best Local Similarity 54.7%; Pred. No. 2.8e-22;  
Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;  
Qy 3635 GCCAGGGAGACGATGCGCGCTTGAGGACGAGATGTCTCTGTCGGTTTATCAGCGTC 3694  
Db |||||  
18652 GACAACGGGAAGATCGGAGCGCTTAAACAAAGTGACGAGTGGGTCAATCATTTGC 18593  
Qy |||||  
3695 GCCTTGCTATAGCTTTCTTCCACAGCATAAACGACCTCTGAAGCTTGGCAGAA 3754  
Db |||||  
18592 CCTTTTGTGGTGACTGCGCCCAAGTAAACACCAAGTTTTCACGATGTTCTTG 18533  
Qy |||||  
3755 AGGACGCTTATATGATAGCTGGTGAACCAATCCAGCAAGGTGGCGTGTGACCCCGCC 3814  
Db |||||  
18532 AGTACTGCAATAACTTTTATCTGTAACCTTTTCCAACTAAATGGCGTGAATGTGCC 18473  
Qy |||||  
3815 TTCCCTCTTCCACCGTCAGATGCTGTTTAAAGGAGAGACTCCCGTGTGCCCCACAG 3874  
Db |||||  
18472 ATACTGCTTCCACGGTAAGCAGATGTTAAGTAGCAATACCCCTTGTCTTGTGCCCATTTG 18413  
Qy |||||  
3875 TCTAAGCACCGCTGAGAAGGAGCGGTGAACCCAGGTACAGTATTTGGCTATCTCTTTTAAAA 3934

Db 18412 ACTAATAACCACTTTGATGGCAATCCGAAATATCTTGTGTGAGTCTTTATAT 18353  
 QY 3935 ATATTTTTCAGCTTTGGAGGTATCTGTAGTCCGAGCAACGCTAAAGGCTAGTCCGGTG 3994  
 Db 18352 ATATTTAATAGGAGGGGAATGCTACTTTCAGGTTTTTACTGAAAAAGCAAGCCGTGC 18293  
 QY 3995 GCTTG---ACCGCATGTAGGATCCTGCGCAAGGATCAGACTTTTAATATCTCCGGT 4051  
 Db 18292 GCTTGGTTGGTCCATGATAGGATCCTGAGCTAAATATACACCTTTTACATCTCTCAAA 18233  
 QY 4052 TCGCAACATACGACCAACACATAATCTATCTATCGTGGATRAATC 4099  
 Db 18232 GCAGTATATTGAATGCGTTAATAACAATCTTCTTGTGGGGATRAATC 18185

RESULT 15  
 US-09-643-990A-1/c  
 ; Sequence 1, Application US/09643990A  
 ; Patent No. 6528289  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robert D. Fleischmann  
 ; Mark D. Adams  
 ; Owen White  
 ; Hamilton O. Smith  
 ; J. Craig Venter  
 ; TITLE OF INVENTION: The Nucleotide sequence of  
 ; the Haemophilus influenzae Rd Genome, Fragments  
 ; Thereof, and Uses Thereof  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville,  
 ; STATE: MD  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3 1/2 inch diskette  
 ; COMPUTER: Dell Pentium  
 ; OPERATING SYSTEM: MS DOS V6.22  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/643,990A  
 ; FILING DATE: 23-Aug-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/487,429  
 ; FILING DATE: 1995-06-07  
 ; APPLICATION NUMBER: 08/426,787  
 ; FILING DATE: 1995-04-21  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kenley K. Hoover  
 ; REGISTRATION NUMBER: 40,302  
 ; REFERENCE/DOCKET NUMBER: PB186P1C1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 301-610-5790  
 ; TELEFAX: 310-309-8439  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1830121 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-643-990A-1

Query Match 1.3%; Score 117.6; DB 4; Length 1830121;  
 Best Local Similarity 54.7%; Pred. No. 2.8e-22;  
 Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;  
 QY 3635 GCCAGGGGAGCGGATCGCGCTTTCAGGACGAGATGCTCTGTGCGTTATCAGGTC 3694

Db 18652 GACACGGGGAGGATGCGGAGCCGTTAAAAACAAAGGTGACGAGTGGGTCAATCATTTGC 18593  
 QY 3695 GCCTTCTCTATAGCCTTTCTTCCCAACAGCAATAAAAAACGACCTCTGAAGCTTGGCAGAA 3754  
 Db 18592 CCTTTTTCGTCGCTGACTGCCCCAAAGTAAAAACACAGTTTTTTCAGATGTTCTATG 18533  
 QY 3755 AGCAGCTTATTTATGTAGCTGGTGAACCAATCCCAAGCAAGGTGGCGTGTGACCCCGCC 3814  
 Db 18532 AGTACTGCAATAACTTTTATCTGTAAACCTTTCCCAACCTAAATTTGGCGTGTGAATGTGCC 18473  
 QY 3815 TTCCCTCTTCCACCCCTCAGATGGTGTTTAAAGCAGAACTCCCGTTTTTGGCCCAACAG 3874  
 Db 18472 ATACCTCGTTTCCACCGGTAAAGCAGAGTGTAAAGTAGCAATACCCCTTGTCTGCCCATTTG 18413  
 QY 3875 TCTAAGCACCCGTGAGAAGGAGCGGTGAACCCAGGTACAGTATTTGGCTATCTCTTTAAAA 3934  
 Db 18412 ACTAATAACCACTTTGATGGCAATCCGAAATATCTTGTGTGAGTCTTTATAT 18353  
 QY 3935 ATATTTTTCAGCTTTGGAGGTATCTGTAGTCCGAGCAACGCTAAAGGCTAGTCCGGTG 3994  
 Db 18352 ATATTTAATAGGAGGGGAATGGGTACTTTCAGGTTTTTACTGAAAAAGCAAGCCGTGC 18293  
 QY 3995 GCTTG---ACCGCATGTAGGATCCTGCGCAAGGATCAGACTTTTAATATCTCCGGT 4051  
 Db 18292 GCTTGGTTGGTCCATGATAGGATCCTGAGCTAAATATACACCTTTTACATCTCTCAAA 18233  
 QY 4052 TCGCAACATACGACCAACACATAATCTATCTATCGTGGATRAATC 4099  
 Db 18232 GCAGTATATTGAATGCGTTAATAACAATCTTCTTGTGGGGATRAATC 18185

Search completed: April 11, 2004, 08:05:57  
 Job time : 471 secs



QY	765	CCACGGGACGCAAAAGACGGAGATATCTGTCGGAGAACATCTATACTTATGATAGG	824
Db	14344	ACACCAAAAACCTGCGATGCTCGTCCGGCAGAGGAAAACTGCTCACTGATCTCATCTG	14285
QY	825	GGAAGAAGTCGTATCTGAGTAGTACACTCTGGAGAGAACTGGCGCTGTTTATAACAA	884
Db	14284	CAACGCGACCTTACCGGAGTACACGGCTCTCGCAGCTGGGCAATTTTCATTACAA	14225
QY	885	CAAGCGCTGCACCGATCTGCAGTTCGGACACTTCTTAAAAATTTTAGATATAATCTTCC	944
Db	14224	CAAGCGATGTGTGGAGCAGAGTTTCGGAACTCATGAAGTGTGAGTAGTACGGCTTCC	14165
QY	945	CAATTCCTGACGAAGTCATGAGTTACGTAGACAGATTGTGCTCTCTAAAAGTAGAGATTAT	1004
Db	14164	CATCACCGAGGACATGCAAGTTGTGGAACGCTTTGCTCCGGAAAGTTACATCAC	14105
QY	1005	GGATCCTTTAGAAATACATTGGCTGACCCGACTCTTTTGTACATACGGAGGTAAAGGC	1064
Db	14104	CAACCGGCAACCTTCGGGGTGACCGCGTGTCTCGTCCACAAGGAGGTACGCGC	14045
QY	1065	GTATCTAACAAACCTACACAATGCTTAACGCTAGGGCGGACATACCGGACACGAAGCT	1124
Db	14044	GTACATGGCCAAAGCTCCACGCCCTAGCTAAAGGTGACTCCGA---GGGGGAGTTTGTGT	13988
QY	1125	CTTTACCTGTCGGTGTGCGAGGTGTTGTGAAGCGTTTGAGGAATACAAAGCGGC	1184
Db	13987	GTTTACCTCCCGTGCTCAGTTTGTGTCGTGCGTCAAGAGTTTGACGAGTATCAGCGCT	13928
QY	1185	CGTCAACCTCACCAACCTCACCGTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAG	1244
Db	13927	CAGCAGCAACCAACCGCTGACCATGGAAGAAGTGATACGGCCCAACGCCATCGCATCAC	13868
QY	1245	TAAATATTCCGAGTTTGTGGACCGAGCATGTCATATCGCCACCGGAATC-----CAC	1298
Db	13867	CAACTACTCCAGAGTCAGGACGAGACCGCGGGCAGCTGCGCTGTGAGGTGCACAGCA	13808
QY	1299	TGAACGTTCGACCCAGGTTACCTTTTACCAAGTTTGTTAAAAACGCCACGCTGTCTTT	1358
Db	13807	GCACAGCTAGTCTGTGGCCCGGAACGACATCACGTACGTCCTCAACGCCAGGTTCGCGT	13748
QY	1359	AAATGGAATAACAAAAATGCAATATGTGGTTTCAGGGTAGCTACTTTGAGTTCAAAG	1418
Db	13747	GACCGCGCGCTCCGAAAGATGGTGTGTTGGGTTCGACGGGAGCTTTCGACCTTCGAGGC	13688
QY	1419	AATCTTAGACAGGAACTCTTCTGTGAGACGCAATTCGAAGACCGCTCGGAATACGTATA	1478
Db	13687	TGTGCTGCGGACGACAGCTTCTGTGAAGACCAAGGGGAGACCTCGGTGGAGTTCGCCTA	13628
QY	1479	CGGTTTTCTTAAACACGCTACTGTACAAACCCATGTACTGTTTCAAC	1524
Db	13627	CCGGTTCCTGTCGGGCTCATGTTCCGGCGGCTGATTCACITTTTAC	13582

RESULT 2  
US-09-913-514-2  
; Sequence 2, Application US/09913514  
; Publication No. US20030082210A1  
; GENERAL INFORMATION:  
; APPLICANT: GOMI, Yasuyuki  
; APPLICANT: SUNAMACHI, Hiroki  
; APPLICANT: TAKAHASHI, Michiaki  
; APPLICANT: YAWANISHI, Koichi  
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccine  
; FILE REFERENCE: 0216-0454P  
; CURRENT APPLICATION NUMBER: US/09/913,514  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: PCT/JP01/00678  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: JP 2000-62734  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Seqin version 3.1.

```

; SEQ ID NO. 2
; LENGTH: 125157
; TYPE: DNA
; ORGANISM: Varicella virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(125157)
; OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

Query Match      1.7%; Score 154; DB 10; Length 125157;
Best local similarity 56.2%; Pred. No. 4.2e-32;
Matches 289; Conservative 0; Mismatches 225; Indels 0; Gaps 0

Qy      585 CAATATTATTGTAATCGACGAGCGCGGAACCCGTGCTCTTACATPACTCACACCGTGT 644
Db      96652 CAACTGTTATGTTATGATGAGCAGGATTGCTAGGGCGTCATATCTTCACGGCGGTGT 96711

Qy      645 GTTCTTTTACTGGTTCCTTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGCGGT 704
Db      96712 TTACTCTGGTGGCTTTTGAATGCTATATATCAAAAGCCCTCAGTACATAAACGGTCGAA 96771

Qy      705 TCCGTGCATAGTATGGTCGGTTCGCCAACGCAGACAAACGGGTCTCCAGTCAACTTACAA 764
Db      96772 ACCGGTCATAGTATGGTCGGTTCGCCAACGCCAACTGACTGCTGTAGAACTCAITTTCA 96831

Qy      765 CCACGGGAGCGCAAAAGACGGAGATATCGTGTGCGAGAACATCTCTTAAACATTCATCATAGG 824
Db      96832 ACATGACATGCAGCGTTCACACGTAACTCCTAGTGAAAATATACTCACGTATATACTGT 96891

Qy      825 GAAGAAGTTCGTATCTGAGTACGTACACTTGGAGAGAACTGGCGCTGTTTTATAAACAA 884
Db      96892 CAATCAAACTCTGCGTCATATACTAAACATCTCAATACTGGGCAATCTTTATTATTA 96951

Qy      885 CAAGCGCTGCACCGAATCTGCAGTTCGGACACCTTTCTAAAAATTTTAGAGTATAAATCTCC 944
Db      96952 CAAACGATGTCAAGAGGACGATTTTGGAAATCTTTTAAAAACGCTTAGTACGGGTACC 97011

Qy      945 CATTCCTGACGAAGTCATGAGTTACGTAGACAGATTGTTCGTTTCTTAAAGTAAAGTAT 1004
Db      97012 TATTACCGAAGCACATGCGCGTCTGGTCGATACATTTGTTGACCTGCATTCCTTATTA 97071

Qy      1005 GGATCCTTTTAGAATAATTTGGCTGGACCGCACTCTTTTGTCCACATAGCGAGGTAAAGC 1064
Db      97072 CAATCTCTGTAATCTCCCGGATGGACGGCTCTGTATTCTGTCGATTAAGGAGGTAGCGC 97131

Qy      1065 GTATCTAAACAAACCTACACATGCTTAACGCTA 1098
Db      97132 GTATATGATGTAAGTATACAGCGCATTTAAAGCTA 97165

```

### RESULT 3

```

US-09-913-514-1
; Sequence 1, Application US/09913514
; Publication No. US200308210A1
; GENERAL INFORMATION:
; APPLICANT: GOMI, Yasuyuki
; APPLICANT: SUNAMACHI, Hiroki
; APPLICANT: TAKAHASHI, Michiaki
; APPLICANT: YAMANISHI, Koichi
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vacci
; FILE REFERENCE: 0216-0454P
; CURRENT APPLICATION NUMBER: US/09/913,514
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/JF01/00678
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-62734
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 124884
; TYPE: DNA

```

; ORGANISM: Varicella virus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(124884)  
; OTHER INFORMATION: Dumas Strain  
US-09-913-514-1

Query Match 1.7%; Score 152.4; DB 10; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1.2e-31;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 585 CAATATTATTGTAATGACGAGCGCGGAAACCCCTCTCTTATCATACTACACCGCGT 644  
Db 96706 CAACGTTATTGTTTATGATGAAGAGGATGCTAGGCGGTCAATCTCTACGCGCGTGT 96765

QY 645 GTTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGCGGT 704  
Db 96766 TTACTGTGTGGCTTTTGAATGCTATATCAAGGCCCTCAGTACATAAAACGTCGAAA 96825

QY 705 TCCGTGATAGTAGTGGTTCGTCGCAACGACAGCAAAACGCTTCCAGTCAACTTACAA 764  
Db 96826 ACCGGTCATAGTAGTGGTTCGTCGCAACGACAGCAAAACGCTTCCAGTCAACTTACAA 96885

QY 765 CCAGCGGACGCAAAAGACGGAGATATCGTGTGCGAGAAACATCTTAACTATATAGG 824  
Db 96886 ACATGACATGCGGTTTCAACAGTAACTCTCTAGTGAATAATATACTACGTATATACTG 96945

QY 825 GAAGAGGTCGTATCTGAGTACGTACACTTGGAGAGAACTGGCGCTCTTTTATAACAA 884  
Db 96946 CAATCAAACCTCTGCTCAATATACTAACTCTCAATACTGGGCAATCTTTTATAATAA 97005

QY 885 CAACGCTGCAACCATCTGAGTTCGGACACCTCTTAAAAATTTTAGAGTATATCTTCC 944  
Db 97006 CAACGATGTCAGAGACGATTTTGAATCTTTTAAAAACGCTTCTGAGTACGGCTACC 97065

QY 945 CAITTCCTGACGAATCATGAGTACGTAGACAGATTTGTCTCTTAAAGTAGATATAT 1004  
Db 97066 TATTACCGAAGACATCGCGTCTGTCGATACATTTTGTACCTGCATCTATATAA 97125

QY 1005 GGATCCTTTAGATATACATTTGGTGGACCGACTCTTTTGTACATAGCAGGTAAGGC 1064  
Db 97126 CAATCCTGCTAATCTTCCGATGGAGCGCTGTATTCGTCGATAGGAGGTGAGCG 97185

QY 1065 GTATCTTAACAACTACACATGCTTAACGCTA 1098  
Db 97186 GTATATGATGAAGTTACACGCGCACTTTAAACTA 97219

RESULT 5  
US-10-369-493-47194/c  
; Sequence 47194, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 47194  
; LENGTH: 690  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-369-493-47194

Query Match 1.5%; Score 137.6; DB 15; Length 690;  
Best Local Similarity 57.5%; Pred. No. 4.3e-29;  
Matches 267; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 3638 AGGGAGACGGATGGCGGCTTCAGGACGAGATGCTCTCTCGCTTTATCAGCGTCGCC 3697  
Db 572 AGCGGACGAGATGGCGGCTTCAGTACATGAGGCTTCTATCTATATCGCCCT 513

QY 3698 TTGCTATAGCTTCTTCCCGACAGCATAAAACGACCTCTGAAGCTTGGCAGAGG 3757

; ORGANISM: Varicella virus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(124884)  
; OTHER INFORMATION: Dumas Strain  
US-09-913-514-1

Query Match 1.7%; Score 152.4; DB 10; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1.2e-31;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 585 CAATATTATTGTAATGACGAGCGCGGAAACCCCTCTCTTATCATACTACACCGCGT 644  
Db 96706 CAACGTTATTGTTTATGATGAAGAGGATGCTAGGCGGTCAATCTCTACGCGCGTGT 96765

QY 645 GTTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGCGGT 704  
Db 96766 TTACTGTGTGGCTTTTGAATGCTATATCAAGGCCCTCAGTACATAAAACGTCGAAA 96825

QY 705 TCCGTGATAGTAGTGGTTCGTCGCAACGACAGCAAAACGCTTCCAGTCAACTTACAA 764  
Db 96826 ACCGGTCATAGTAGTGGTTCGTCGCAACGACAGCAAAACGCTTCCAGTCAACTTACAA 96885

QY 765 CCAGCGGACGCAAAAGACGGAGATATCGTGTGCGAGAAACATCTTAACTATATAGG 824  
Db 96886 ACATGACATGCGGTTTCAACAGTAACTCTCTAGTGAATAATATACTACGTATATACTG 96945

QY 825 GAAGAGGTCGTATCTGAGTACGTACACTTGGAGAGAACTGGCGCTCTTTTATAACAA 884  
Db 96946 CAATCAAACCTCTGCTCAATATACTAACTCTCAATACTGGGCAATCTTTTATAATAA 97005

QY 885 CAACGCTGCAACCATCTGAGTTCGGACACCTCTTAAAAATTTTAGAGTATATCTTCC 944  
Db 97006 CAACGATGTCAGAGACGATTTTGAATCTTTTAAAAACGCTTCTGAGTACGGCTACC 97065

QY 945 CAITTCCTGACGAATCATGAGTACGTAGACAGATTTGTCTCTTAAAGTAGATATAT 1004  
Db 97066 TATTACCGAAGACATCGCGTCTGTCGATACATTTTGTACCTGCATCTATATAA 97125

QY 1005 GGATCCTTTAGATATACATTTGGTGGACCGACTCTTTTGTACATAGCAGGTAAGGC 1064  
Db 97126 CAATCCTGCTAATCTTCCGATGGAGCGCTGTATTCGTCGATAGGAGGTGAGCG 97185

QY 1065 GTATCTTAACAACTACACATGCTTAACGCTA 1098  
Db 97186 GTATATGATGAAGTTACACGCGCACTTTAAACTA 97219

RESULT 5  
US-10-369-493-47194/c  
; Sequence 47194, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 47194  
; LENGTH: 690  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-369-493-47194

Query Match 1.5%; Score 137.6; DB 15; Length 690;  
Best Local Similarity 57.5%; Pred. No. 4.3e-29;  
Matches 267; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 3638 AGGGAGACGGATGGCGGCTTCAGGACGAGATGCTCTCTCGCTTTATCAGCGTCGCC 3697  
Db 572 AGCGGACGAGATGGCGGCTTCAGTACATGAGGCTTCTATCTATATCGCCCT 513

QY 3698 TTGCTATAGCTTCTTCCCGACAGCATAAAACGACCTCTGAAGCTTGGCAGAGG 3757

Db 512 TTCTTTGGCGATGCGATCCCAACAAAAACACACGCGCTTCGGGATGCTGGTTAATC 453  
Qy 3758 CAGCTTATTATGATGTTGTAACCAATCCAGCCAGAGTTGGCGTGTACCCCGCTTC 3817  
Db 452 AGGCTGATCACTTTATCGGTGAAGTTTCCAGCCGAGGCTGGCGTGGGAATGCGCCTGA 393  
Qy 3818 CCTCTTTCCACCGTCAGAAATGGTGTAAAGCAGAACTCCCGGTTTTCGCCAACAGTCT 3877  
Db 392 CCTGGCGGTACCGTCAACACAGTATTAGTAGCAGAACGCGCTGACGCGCCAGCTTTC 333  
Qy 3878 AAGCACCCTGAGAACGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAATA 3937  
Db 332 AGATAACCATGATTAGGCGCGGTGAAGCCCGGAATAGTATTTCCAGCTCTTTATACA 273  
Qy 3938 TTTTGGAGGCTTGGAGGTATCTAGTCCGAGCAACGCTAAAGGCTAGTCCGGTGGCT 3997  
Db 272 TTCAATAACGAGGAGAAATGCAATGCGGGGACGAAATGCCAGACATCGGC 213  
Qy 3998 TGACC---GGATGGTAAGATCTGCGCAAGGATCACGCTTTAATATCCTCCGGTTCG 4054  
Db 212 TGTCCCGGTCGTTAAGGATCTGCGCGAGAAATCACCACTTTAAACGTCAOCCAACTCT 153  
Qy 4055 CAACAAATACGACCAACCAATAATCTATCTATCTATCGGTGGATAAAT 4098  
Db 152 GTAAAGCGGACGCGTTAAAGACATCTTTTGTGGGTAGAT 109

RESULT 6  
US-10-260-877-13/c  
; Sequence 13, Application US/10260877  
; Publication No. US20030021813A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Chovan, Linda E.  
; APPLICANT: Hessler, Paul E.  
; APPLICANT: Reich, Karl A.  
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME  
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF  
; TITLE OF INVENTION: 'ESSENTIAL GENES'  
; FILE REFERENCE: 6565.US.F1  
; CURRENT APPLICATION NUMBER: US/10/260,877  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: US/09/649,145  
; PRIOR FILING DATE: 2000-08-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 657  
; TYPE: DNA  
; ORGANISM: H. influenzae  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(657)  
; OTHER INFORMATION: HL-0018  
US-10-260-877-13

Query Match 1.3%; Score 117.6; DB 14; Length 657;  
Best Local Similarity 54.7%; Pred. No. 3.6e-23;  
Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

Qy 3635 GCGAGGGGAGCGGATGCGGCGCTTGAGGACGAGATGCTCTGTCGTTTATCAGCGTC 3694  
Db 566 GACAAACGGGAGGATGCGGAGCGCTTAAACAAAGGTGACGAGTGGCGGTCAATCATTTGC 507  
Qy 3695 GCCTTGCTATAGCTTTCTTCCCAACAGATATAAAACGCACTCTGAAGCTTGGCAGAA 3754  
Db 506 CCTTTTGTGGGTGACTGCCCCCAAGTAAACACCAAGTTTTCACGATGTTTATTG 447  
Qy 3755 AGGACGCTTATTATGATAGCTGGTGAACCAATCCAGCGCAAGTTGGCGTGTGACCCGCC 3814  
Db 446 AGTACTGCAATAAATTTATCTGTAAACCTTTCCCAACCTAAATTTGGCGTGTGAATGTC 387  
Qy 3815 TTCCCTCTTTCCACCGTCAGAAATGCTGTTTAAAGCAGAACTCCCGCTTTTGCCCAACAG 3874

Db 386 ATACTCGTTCACCGTAAGCACAGTGTAAAGTAGCAATACCCCTTGTCTGCCCATTTG 327  
Qy 3875 TCTAAGCACCGCTGAGAGGAGCGGTGAACCCAGGTACAGTATGCTCTTTAAAA 3934  
Db 326 ACTAATAACCATTTGATGGCAATTTGAATCCCGAATACTTGTGTGAGTCTTTATAT 267  
Qy 3935 ATATTTTGGAGTTGGAGGTATCTAGTATGTCGAGCAACGCTAAAGGCTAGTCCGGTG 3994  
Db 266 ATATTAATAGGAGGAGGGAATGGCTACTTACGTTTCTGAAAAAGCCAGCCGTC 207  
Qy 3995 GCTTG---ACGCGCATGTAAGGATCTGGCCAAAGGATCAGACTTTTAATATCTCCGGT 4051  
Db 206 GCTTGGTTGGTCCATGATAAGGATCTGACCTAAATACCCTTTTACATCTCTCAAA 147  
Qy 4052 TCGAACAATAACGACCAACCAATAATTTCTATCTATCGGTGGATAAATC 4099  
Db 146 GCGATATATTGAATCGGTAAATACATCTTCTTGTGGGGATAAATC 99

RESULT 7  
US-10-329-960-1/c  
; Sequence 1, Application US/10329960  
; Publication No. US20030099277A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra  
; TITLE OF INVENTION: Thereof, and Uses Thereof  
; FILE REFERENCE: PB186P1  
; CURRENT APPLICATION NUMBER: US/10/329,960  
; CURRENT FILING DATE: 2003-01-02  
; PRIOR APPLICATION NUMBER: US 09/643,990  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: US 08/487,429  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: US 08/426,787  
; PRIOR FILING DATE: 1995-04-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1830121  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4747)..(4747)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (9921)..(9921)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (10150)..(10150)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (29238)..(29298)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36543)..(36543)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36551)..(36551)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36636)..(36636)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature



```
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (133360)..(133360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match      1.3%; Score 117.6; DB 14; Length 1830121;
Best Local Similarity 54.7%; Pred. No. 2.4e-20;
Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 3635 GCACGGGAGAGAGGAGGCGCGCTTGAGGAGAGAGATGCTGTCGCGTTATCAGCGTC 3694
Db 18652 GACACGGGAGAGATGCGGAGCGGTAAACAGGTGACGAGTGGCGTCAATTTGCG 18593

QY 3695 GCCTTCTATAGCTTTCTTCCACACAGCATAAAACGCACTCTGAGCTTGGCAGAA 3754
Db 18592 CCTTTTGTGCGTGACTGCCCCAAAGTAAACACACAGTTTTCAGATGTTCAATG 18553

QY 3755 AGCAGCTTATTATGATGCTGGTGAACCAATCCAGCCAGGTGCGGTGACCCCGCC 3814
Db 18532 AGTACTGCAATACTTTATCTGTAAACCTTTCACACCTAAATTTGCGGTGAATGTGC 18473

QY 3815 TTCCCTCTTCCACCGTCAGATGGTGTAAAGCAGACTCCCGTTTTCGCCACAG 3874
Db 18472 ATACCTCGTTCCACGGTAAGCAGTGTAAAGTAGCAATACCCCTTGTTCGCCATTTG 18413

QY 3875 TCTAAGCACCGGTGAGAGGAGCGGTGAACCCAGGTACAGTATGCTATCTTTAAAA 3934
Db 18412 ACTAATAACCATTTGATGCAATTTGAATCCCGAAATATCTTGTGAGTCTTTATAT 18353

QY 3935 ATATTTTGGCTTGAGGTATCTGTAGTCCGGAGACGCTAAAGCTAGTCCGGTG 3994
Db 18352 ATATTTAATAGGAAGGGAATGGCTACTTCCAGTGTGTTTACTGAAAAGCCAGCGTGC 18293

QY 3995 GCTTG---ACCGGATGTAAGGATCCTGCCAAGGATCAGGACTTAAATATCCTCCGCT 4051
Db 18292 GCTTGGTTGGTCCATGATGAAGATCTGACCTAAATACCACTTTTCACTCCACAAA 18233

QY 4052 TCGCAACAATACGACCAACCATATTTCTATCTATCGGTGGAATATC 4099
Db 18232 GCAGTATATTGATGCGTTAAATACATCTTCTGTGGGGGATAAATC 18195

```

```

RESULT 8
US-10-329-670-1/c
Sequence 1, Application US/10329670
Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
FILE REFERENCE: P186P1
CURRENT APPLICATION NUMBER: US/10329.670
CURRENT FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

```

```

; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)

```

```

; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match      1.3%; Score 117.6; DB 15; Length 1830121;
Best Local Similarity 54.7%; Pred. No. 2.4e-20;
Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 3635 GCACGGGAGACGGATGGCGCCCTTGAGGACGAGATGTCTGTCCGTTTATCAGCTC 3694
Db 18652 GACACGGGGAAGGATGCGGAGCCGTTAAACAAGTGACGAGTCGGTCAATCTTGC 18593
QY 3695 GCCTTGTCTATAGCCCTTCTCCCCACAGCATATAAAGCCACCTCTGAAGCTTGGCAGAA 3754
Db 18592 CCTTTTTTTTGGGTGACTGCCCCCAAGTAAACACCCAGTTTTTACGATGTTCAATG 18533
QY 3755 AGCAGCTTATATGTAGCTGTGAACCAATCCAGCCAGGTTGGCGGTGTGACCCGCC 3814
Db 18532 AGTACTGCAATAACTTTATCTGTAAACCTTTCCCAACCTAAATGGCGGTGGAATGTGCC 18473
QY 3815 TTCCCTCTTTCCACCGTCAGATGGTGTAAAGACAGAACTCCCGCTTTTGCCCAACAG 3874
Db 18472 ATACCTCGTTCCACGGTAAGCAGCTGTAAGTASCAATACCCCTTGTCTGCCCATTTG 18413
QY 3875 TCTAGCACCCGTGAGAGGACGGTGACCCAGTACAGTATGGCTATCTCTTTAAAA 3934
Db 18412 ACTAAATAACCATTTGATGGCAATTTGAAATCCCGAAATATCTGTGTAGTCTTTATAT 18353
QY 3935 ATATTTTGTAGGCTTGGAGGTATAGTGTAGTCCGAGCAAGCTAAAGGCTAGTCCGGTG 3994
Db 18352 ATATTTAATAGGAAGGGGATGGCTACTTCAGSTTTTACTGAAAAGCAAGCGGTGC 18293
QY 3995 GCTTG---ACCGGATGTAGGATCCTGGCAGGATCAGACTTTAATATCTCTCCCGT 4051
Db 18292 GCTTGTTTGGTCCATGATAAGGATCCTGACCTAAATATACCACTTTTACATCTCTCAAA 18233
QY 4052 TCGCAACATACGACCAACCATATTTCTATCTATCGGTGGATAAATC 4099
Db 18232 GCAGTATATTTGAATGCGTTAATACATCTTCTTGTGGGGGATAAATC 18185

```

## RESULT 9

```

US-10-369-493-44794/c
; Sequence 44794, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

```

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 44794  
 ; LENGTH: 675  
 ; TYPE: DNA  
 ; ORGANISM: Xenorhabdus nematophilus  
 US-10-369-493-44794

Query Match 1.2%; Score 110.6; DB 15; Length 675;  
 Best Local Similarity 53.5%; Pred. No. 4.3e-21;  
 Matches 255; Conservative 0; Mismatches 219; Indels 3; Gaps 1;  
 QY 3635 GCCAGGAGAGGATGCGCGCTTGGAGCAGAGATGCTCTGTCGGTTATTCACGGTC 3694  
 DB 575 GACAAAGAGAGGATGCGCGCTTGGAGCAGAGATGCTCTGTCGGTTATTCACGGTC 516  
 QY 3695 GCCTTGCTATAGCCTTCTTCCCAACAGCATAAACCGACCTCTGAAGCTTGGCAGAA 3754  
 DB 515 CTTTCTTCTGGCATGGGACCCCAAGTAGGAATAGCGCGTGACGATGTTTCATG 456  
 QY 3755 AGGACGCTATATGAGTGGTGAACCAATCCAGCAGAGTGGGGTGAACCGCC 3814  
 DB 455 ATTGCCGCAATGACTTGTGCTAGTGAAGTTTCCAGCCTAAATTCGCATGTGAATGCGCA 396  
 QY 3815 TTCCCTCTTCCACCGTCAGATGGTGTGTTAAAGCAGAACTCCCGGTTTGGCCCAACAG 3874  
 DB 395 TTGCCACGTTCAACGGTCAGCACCCTATTCAGCAGCAATAGCCCTGCTTGGCCCACTG 336  
 QY 3875 TCTAAGACCGCTGAGAGAGGCGGTGAACCCAGCTACATATGCTATCTCTTTAA 3934  
 DB 335 ACCAAGCAGCATGATTAGGCGCGCTTAAACCCAGTAAATTCGATTCACGCTCTTTATAC 276  
 QY 3935 ATATTTTGGGCTTGGAGGTATCTGATGTCGGAGCAACGCTAAAGGCTAGTCCGGTG 3994  
 DB 275 ATGTTGACAGAGAGTGGTGCAGGAATACAGGCTGATCTGAGAAAGCCAGCGTGT 216  
 QY 3995 GCTTCA---CGCGATGTAAGATCCTGGCAGAGATCAGCTTAAATCTCCGCT 4051  
 DB 215 GCCTGATTGGCGCGTGGTGAAGGCTTTGGCCCAAAATCAGACTCTTACATCGGACGT 156  
 QY 4052 TCGCAACAATACGACCAACCATATTTCTATCTATCGGTGGATAAATCAGGTTAGC 4108  
 DB 155 TGGTATAGCGAATGCAATGGAACATCTCTGCTGTAGCGGATAAATGCTTTCCG 99

RESULT 10  
 US-09-758-017A-1/c  
 ; Sequence 1, Application US/09758017A  
 ; Patent No. US2002015557A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lanes, Olav  
 ; APPLICANT: Willasen, Nils Peder  
 ; APPLICANT: Guddal, Per Henrik  
 ; APPLICANT: Gjellesvik, Dag Rune  
 ; TITLE OF INVENTION: Cod uracil-DNA glycosylase, gene coding therefore,  
 ; TITLE OF INVENTION: recombinant DNA containing said gene or operative parts  
 ; TITLE OF INVENTION: thereof, a method for preparing said protein and the  
 ; TITLE OF INVENTION: use of said protein or said operative pa  
 ; FILE REFERENCE: U013209-3  
 ; CURRENT APPLICATION NUMBER: US/09/758,017A  
 ; CURRENT FILING DATE: 2001-01-10  
 ; PRIOR APPLICATION NUMBER: 2000-5428  
 ; PRIOR FILING DATE: 2000-10-27  
 ; PRIOR APPLICATION NUMBER: 2000-0163  
 ; PRIOR FILING DATE: 2000-01-12  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1

; LENGTH: 1283  
 ; TYPE: DNA  
 ; ORGANISM: Gadus morhua  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (18)..(920)  
 US-09-758-017A-1  
 Query Match 1.2%; Score 107.4; DB 9; Length 1283;  
 Best Local Similarity 51.4%; Pred. No. 6.6e-20;  
 Matches 274; Conservative 0; Mismatches 256; Indels 3; Gaps 1;  
 QY 3632 GTAGCCAGGGAGAGCGATGCGCGCTTGGAGCAGAGATGCTCTGTCGGTTATTCACG 3691  
 DB 829 GCAGACAAAGAGATGGAACAGCTTGCACAGACATGGTGAGCTTCTCTGTCGATGTC 770  
 QY 3692 GTGCGCTTGTCTATAGCTTCTTCCCAACAGCATAAACCGACCTCTGAAGCTTGGCA 3751  
 DB 769 GCTCCCTTCTTATGCGGCTATGAGCCCCCAACAGGAAACGACTCTTCCCGTTGACG 710  
 QY 3752 GAAAGGAGCTTATATGATGCTGTGAACCAATCCAGCAGAGTGGCGTGTGACCCC 3811  
 DB 709 CTCAGCCACTGTATCAGCGTCGGTGAAGTCTCCAGCCTCTGCTCTTGGGAGTTG 650  
 QY 3812 GCCTTCCCTTTCACCGCTCAGATGGTGTGTTAAAGCAGAACTCCCGGTTTGGCCAA 3871  
 DB 649 GCCTGATGGCCCGCAGCTGACGACCGCTTAAAGCAGCAGCACCCTTGTGTTGCCAT 590  
 QY 3872 CAGTCTAAGCACCCTGGAAGAGCGGTGAACCCAGGTACAGTATGCTATCTCTTTA 3931  
 DB 589 CCGCTTAGATCTCCATGTCAGGATGCTTGAAGCCATCAATGCGGTACACAATCTTTG 530  
 QY 3932 AATAATTTTGGAGCTTGGAGTACTGTAGTCCGAGCAGCTTAAGCTAGCTAGCTCCG 3991  
 DB 529 TATATGTCAGAGACTGGGGGAGGGGAACTGGCTTTTGACACTGAAACAGAGTCCA 470  
 QY 3992 GTGCTTGACCGGA---TGTTAAGGATCCTGGCCAAAGATCAGCACTTTAATATCTCTC 4048  
 DB 469 TGTGCTTGGTGGACCGTGTGAAGGCTCTGCGCTTGAATCACTACTTTTCACTTTGA 410  
 QY 4049 GGTTCGCAACAATACGACCAACATATCTATCTATCGTGGTAATCAATCAGGTTAGC 4108  
 DB 409 ATGTACACATCTCTGTCGAACCTGTACACTTGTATCAGCCGGTGGTGAAGCTGACGG 350  
 QY 4109 TTGCTCATATCCATAACGGCTTTCAGCAGAGCGCGAGCTTCTGTTTAAAAA 4161  
 DB 349 CTCCTCTCATCAGCTACAAAGGACATCAATGTTTGAAGTATGGCTTTTCAAA 297

RESULT 11  
 US-10-369-493-24211/c  
 ; Sequence 24211, Application US/10369493  
 ; Publication No. US2003233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 24211  
 ; LENGTH: 744  
 ; TYPE: DNA  
 ; ORGANISM: Deinococcus radiodurans  
 US-10-369-493-24211

Query Match 1.2%; Score 106; DB 15; Length 744;  
Best Local Similarity 55.4%; Pred. No. 1.1e-19;  
Matches 227; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

Qy 3638 AGGGAGAGCGATGCGCGCTTGGAGGACGAGATCTCTCTGTCGGTTTATCAGCGTGGCC 3697  
Db |||||  
Qy 629 AGCGGAGGAGTCCCGACTCGATGACGAGCTGTTTGGCGGTAATCAGCTTCTTT 570  
Db |||||  
Qy 3698 TTGTTCTATAGCTTCTTCCACAGCATATAAAGCACCTCTGAAGCTTGGCAGAAGG 3757  
Db |||||  
Qy 569 TTCTTGGCGGCTAGTCTCCCGAGAGGATAAAGACGACGCTCTCTTGGCGTGGACG 510  
Db |||||  
Qy 3758 CAGCTTATATAGTCTGGTGAACCAATCCAGCCAGGTTGGGCTGTGACCCCGCTTC 3817  
Db |||||  
Qy 509 GCCTTGATCAGCGCTGGTGAAGTGTCCAGGCTTTGGCTGTGGCTGTGGCTGC 450  
Db |||||  
Qy 3818 CTTCTTTCCAGCTCAGTAATGTGTATAAAGCAGAACTCCCGCTTTTGGCCACAGTCT 3877  
Db |||||  
Qy 449 CCGCGCCGCGCTGAGCAGCGCATTTGAGCAGCAAGACGCGCTGCTCGGCCCATGAACGC 390  
Db |||||  
Qy 3878 RAGCACCGGTGAGAGGCGGTGAACCCAGCTACAGTATTGGCTATCTCTTTAAATA 3937  
Db |||||  
Qy 389 AGATACCGCTGCTGGCGCCACGAGCGGGGATGCTCTGCTGAGTCTCTTGTAGATG 330  
Db |||||  
Qy 3938 TTTTGAAGCTTGGAGTATATCTAGTCCGAGCAAGCTTAAAGGTAGTCCCGTGGCT 3997  
Db |||||  
Qy 329 TTGCGAGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 270  
Db |||||  
Qy 3998 TGA---CGCGATGTAGGATCTCTGGCGCAAGGATCAGACTTTAATATC 4044  
Db |||||  
Qy 269 TGATTGGGCGCTGTAGGGTCTCGCGCGAGATCAGACCTTGACCTC 220  
Db |||||

RESULT 12  
US-10-369-493-23923/c  
; Sequence 23923, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; NUMBER OF SEQ ID NOS: 2002-02-21  
; SEQ ID NO 23923  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Xenorhabdus nematophilus  
US-10-369-493-23923

Query Match 1.2%; Score 105.8; DB 15; Length 675;  
Best Local Similarity 53.4%; Pred. No. 1.2e-19;  
Matches 246; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

Qy 3641 GGAGCGATGGCGGCTTGGAGGACGAGATGCTCTGTCGGTTTATCAGCGTGGCTTG 3700  
Db |||||  
Qy 569 GGAGAGGATGGTGTCTTTCAGGACATATGGCGCTGGGTATCATTAATAGGCTTTT 510  
Db |||||  
Qy 3701 TCTATAGCTTTCTTCCACAGCATATAAAGCGACCTCTGAAGCTTGGCAGAAGGCAG 3760  
Db |||||  
Qy 509 TTTCTGAGCATGACGACCCAGAGGAGGAAATAAACACACTTCGGTGTTCATTGATCGCA 450  
Db |||||  
Qy 3761 CTTATATAGTGTGGTGAACCAATCCAGCCAGGTTGGCTGTGACCCCGCTTCCCT 3820  
Db |||||  
Qy 449 GCAATGACCTTATCGGTGAAGTCTCCAGCCAGATTTGCATGAGATGAGCATTACCC 390  
Db |||||

Qy 3821 CTTTCCACCGTCAGATGGTGTTRAAAGCAGAACTCCCGTTTTCGCCCAACAGTCTAAG 3880  
Db |||||  
Qy 389 CGTCAACTGTTHACACTGTATTTCAGTATRAAACCCCTGTGTTTGCCACAGCTAAGCAG 330  
Db |||||  
Qy 3881 CACCCGTGAGAAAGACCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAATAATTT 3940  
Db |||||  
Qy 329 CAGCCGTGATTTGGGACGCTGMAATCTGCAATATCAGATTCAGATTCCTTTGTACATATTA 270  
Db |||||  
Qy 3941 TTGAGGCTTGGAGGTATCTGTAGTCCGGAGCAACCTTAAAGGCTAGTCCGGTGGCTTGA 4000  
Db |||||  
Qy 269 ACAGGTGAAGGTGGAGCGGGAATGCCCGCTGTACTGAAAGACGAGGCATGAGGCTGA 210  
Db |||||  
Qy 4001 ---CCGCGATGTGAAGATCTCTGGCCCAAGGATCAGACTTTAATATCTCCGTTTCGCA 4057  
Db |||||  
Qy 209 TTCCGGCGGTGATAAGGGTCTTGTCCAGATCAGACTTCACATCAGCCAGTTCGGTA 150  
Db |||||  
Qy 4058 CAATGACACCAACATATTTCTATCTATCGGTGGATAAAT 4098  
Db |||||  
Qy 149 TAGCGAAAGCGTTGAAAACATCTCTGTGGCGGATAAAT 109  
Db |||||

RESULT 13  
US-10-369-493-28257/c  
; Sequence 28257, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; NUMBER OF SEQ ID NOS: 2002-02-21  
; SEQ ID NO 28257  
; LENGTH: 594  
; TYPE: DNA  
; ORGANISM: Burkholderia fungorum  
US-10-369-493-28257

Query Match 1.2%; Score 104.6; DB 15; Length 594;  
Best Local Similarity 56.6%; Pred. No. 2.4e-19;  
Matches 244; Conservative 0; Mismatches 169; Indels 18; Gaps 2;

Qy 3635 GCCAGGGGAGCGATGGCGGCTTGGAGGACGAGATGCTCTGTCGGTTTATCAGCGTC 3694  
Db |||||  
Qy 506 GACAGCGGCGAGGATGGCGTCTTCAGACGCAATGCACTTGGCGCGGAGCGCGC 447  
Db |||||  
Qy 3695 GCCTTGTCTATAGCTTTCTTCCCAAGCAATAAACCACCTCTGAAGCTTGGCAGAA 3754  
Db |||||  
Qy 446 CGCTTGGCTTGGCATGGCGGCCCAACAGCATGAACAGGCGCTCATGGCGCATCGCG 387  
Db |||||  
Qy 3755 AGCAGCTTATATGTAGCTGTGTAACCAATCCAGCCAGGTTGGCTGTGACCCGCGC 3814  
Db |||||  
Qy 386 AGTTCGTGATCAGCGGTGGTGCATTTTCCAGCGCGGTTTGGCTGGCTGGCGGCC 327  
Db |||||  
Qy 3815 TTCCCTCTTTTCCACCGCTCAGATGGTGTTTAAAGCAGAACTCCCGCTTTTGGCCACAG 3874  
Db |||||  
Qy 326 GAGTCTCGCTCCACCGCTCAACCGGTGTTCCAGCAGCAGCAGCGCTCTTCGCCAGGTG 267  
Db |||||  
Qy 3875 TCTAAGCACCGGTGAGAGGCGGTGAACCCAGGTACAGTATGTGCTATCTTTTAAAA 3934  
Db |||||  
Qy 266 TCAGAGACAGCGGTGACGCGCGCTCTCGTGGCCA---AGGCTTGGCGGATCTCTCTTGAAG 210  
Db |||||  
Qy 3935 ATATTTTGGAGCTTGGAGGTATCTGTAGTCCGAGCAACGCTTAAAGGTAGTCCGGTG 3994  
Db |||||  
Qy 209 ATGTTGGCAGCAGCGGCGGCTGGCCAGCTTCCGCGCTACCGAAACGCGCCCGCTGC 150  
Db |||||

QY 3995 GCTTG-----ACCGCATGTAAGGATCTGCCAGGATCAGACTTTA 4039  
Db 149 GCTTGGCGGTCGCCAGGCTTCGCGTGATACGGATCTGCCAGGATCACCACCTTC 90  
QY 4040 ATATCTCCGG 4050  
Db 89 ACTTCATCGG 79

## RESULT 14

US-10-369-493-31014/c  
; Sequence 31014, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 31014  
; LENGTH: 987  
; TYPE: DNA  
; ORGANISM: Burkholderia cepacia  
US-10-369-493-31014

Query Match 1.2%; Score 104.6; DB 15; Length 987;  
Best Local Similarity 56.6%; Pred. No. 3.6e-19;  
Matches 244; Conservative 0; Mismatches 169; Indels 18; Gaps 2;  
QY 3635 GCACGGGAGACGGATGGCGGCTTGAGACAGATGCTCTGTCCGTTATCAGCGTC 3694  
Db 902 GACAGGGGACGGATGGCGGTGTCTCCAGACCAATGGACTTGCCTCCGAGCAGCGG 843  
QY 3695 GCCTTGCTATAGCTTTCTTCCACACAGCATATAAAACGACCTCTGAAGCTTGGCAGAA 3754  
Db 842 CGCTTGCTTGCGATGCGGCGCCACAGCATGAACAGGCGCGTCATGGCGCATCGG 783  
QY 3755 AGGAGCTTATATGATGCTGTGAACCAATCCAGCAGGTTGCGGTGACCCCGCC 3814  
Db 782 AGTTCGTGATCAGCGTGTGCGTGCATTTTCCAGCGCGGTTTTCGCGTGGCGGCC 723  
QY 3815 TTCCCTCTTCCACCGTCAGATGGTGTATAAAGCAGAACTCCCGGTTTGCCTCAACAG 3874  
Db 722 GAGTCTGCTCCACCGTCACACCGTGTTCAGCAGCAGCAGCGCTTCCGCGCAGGTG 663  
QY 3875 TCTAAGCACCGGTGAGAGAGCGGTGAACCCAGGTACAGTATTTGGCTATCTCTTTAAA 3934  
Db 662 TCGAGACAGCGGTGACCGCGCTCTCGTGCCCA---AGGCTTGGCGCGATCTCTTTGAAG 606  
QY 3935 ATATTTTGGCTTGAGGTATCTAGTCCGAGCAGCAGCTAAAGGCTAGTCCGCGTG 3994  
Db 605 ATGTTCGCGAGCAGCGCGCGTGCAGCGCTTCGCGCTTACCGAAACCGCGCGCGTGC 546  
QY 3995 GCTTG-----ACCGCATGTAAGGATCTTCGCGCAAGGATCAGACTTTA 4039  
Db 545 GCTTGGCGGTGCCACGGTCTTCGCGTGATACGGATCTTCGCGCAGGATCACCACCTTC 486  
QY 4040 ATATCTCCGG 4050  
Db 485 ACTTCATCGG 475

## RESULT 15

US-10-369-493-32380/c  
; Sequence 32380, Application US/10369493

; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 32380  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Ralstonia metallidurans  
US-10-369-493-32380

Query Match 1.1%; Score 101; DB 15; Length 675;  
Best Local Similarity 54.8%; Pred. No. 3.1e-18;  
Matches 200; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 3635 GCACGGGAGACGGATGGCGGCTTGAGACAGATGCTCTGTCCGTTATCAGCGTC 3694  
Db 575 GACAGCGGACGGATGGCGGCTTCCAGACGCAATGCGTCTGGCCTTCGAGCAAGCG 516  
QY 3695 GCCTTGCTATAGCTTTCTTCCACACAGCATATAAAACGACCTCTGAAGCTTGGCAGAA 3754  
Db 515 CGCTTGCGCTGCGGTGCTGCCAGAGCAGGATTCGGCCGCGATGCCGCC 456  
QY 3755 AGGAGCTTATATGATGCTGTGAACCAATCCAGCAGGTTGGCGTGTGACCCCGCC 3814  
Db 455 AGCGGTGATCAGCAGTGGTCCAGCTTCCAGCGCGCTTCGATGACTGGCGGCC 396  
QY 3815 TTCCCTCTTCCACCGTCAGATGGTGTATAAAGCAGAACTCCCGGTTTGCCTCAACAG 3874  
Db 395 TGTCCCTGTGACCGCTCAGTACCGTGTTCAGCAGCAGCAGCGCTGACGTGCCAGCT 336  
QY 3875 TCTAAGCACCGGTGAGAGCGGTGAACCCAGGTACAGTATTTGGCTATCTCTTTAAA 3934  
Db 335 TCCAGATTGCCGACGACGCGTCCGCGCCACGCGCTATTCCGCGCGGATCTCTCTGAAG 276  
QY 3935 ATATTTTGGCGTGGAGGTATCTAGTCCGAGCAGCAGCTAAAGGCTAGTCCCGTG 3994  
Db 275 ATGTTCGCGAGGCTTGGTGGACCTTGACGCCCTTCGGGTACGAGAGCGCGCGTGT 216  
QY 3995 GCTTG 3999  
Db 215 GCCTG 211

Search completed: April 11, 2004, 13:37:24  
Job time : 2015 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 20:47:25 ; Search time 13557 Seconds

(without alignments)  
19985.215 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038

Perfect score: 9073

Sequence: 1 atggagagttccgtcgatg.....ggggccacggatgcgcttga 9073

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estl:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estom:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrc:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241.6	2.7	717	29	AG101978 Pan trogl
2	233	2.6	678	29	AG104414 Pan trogl
3	213	2.3	688	29	AG126121 Pan trogl
C 4	194.4	2.1	721	29	AG051826 Pan trogl

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	135	1.5	685	29	AG124918	AG124918 Pan trogl
6	111	1.2	656	29	AG051825	AG051825 Pan trogl
7	105.8	1.2	696	9	AJ395572	AJ395572 AJ395572
8	103.6	1.1	911	13	EU444447	EU444447 603213237
9	103.6	1.1	987	13	EX706705	EX706705 EX706705
10	103.2	1.1	871	13	EX719507	EX719507 EX719507
11	103	1.1	645	9	AL894178	AL894178 AL894178
12	102.4	1.1	699	14	CF521628	CF521628 AGENCOURT
13	101	1.1	806	9	AL669156	AL669156 AL669156
14	100.8	1.1	574	12	BI940552	BI940552 db71f06.Y
15	100.8	1.1	687	14	CD321504	CD321504 StrPUS38
16	100.8	1.1	935	14	CA983550	CA983550 AGENCOURT
17	100.2	1.1	963	13	BU442636	BU442636 603210192
18	100	1.1	823	14	CB938158	CB938158 IPCGX13
19	96.6	1.1	658	12	BJ527177	BJ527177 BJ527177
20	94.8	1.0	909	13	EX754661	EX754661 EX754661
21	93.8	1.0	822	13	EX744303	EX744303 EX744303
22	93.4	1.0	840	14	CD643361	CD643361 AGENCOURT
23	92.2	1.0	717	14	CD764451	CD764451 GGSZLB101
24	92	1.0	612	12	BU064298	BU064298 BU064298
25	91.2	1.0	609	9	AU126319	AU126319 AU126319
26	91	1.0	540	14	CF847247	CF847247 pshB042xG
27	90	1.0	822	12	BI823926	BI823926 603039103
28	90	1.0	871	13	BQ228775	BQ228775 AGENCOURT
29	90	1.0	892	13	BU155625	BU155625 AGENCOURT
30	90	1.0	1052	12	EM541301	EM541301 AGENCOURT
31	90	1.0	1175	12	BM471137	BM471137 AGENCOURT
32	89.8	1.0	773	14	CF947286	CF947286 UI-D-GC1-
33	89.4	1.0	962	14	CA973075	CA973075 AGENCOURT
34	88.8	1.0	622	12	BI066406	BI066406 pglfn.pk0
35	88.4	1.0	548	9	AA290918	AA290918 z64b06.r
36	88.2	1.0	871	13	EX423498	EX423498 EX423498
37	88	1.0	1201	13	EX342506	EX342506 EX342506
38	88	1.0	1201	13	EX379137	EX379137 EX379137
39	87.4	1.0	736	12	BU584335	BU584335 BU584335
40	87	1.0	606	13	BQ826724	BQ826724 Gd46b09.X
41	85.6	0.9	786	13	EX887665	EX887665 EX887665
42	84.8	0.9	669	9	AU039323	AU039323 AU039323
43	84.8	0.9	757	12	BU373039	BU373039 BU373039
44	84.2	0.9	913	13	BQ882986	BQ882986 AGENCOURT
45	84	0.9	937	14	CA983409	CA983409 AGENCOURT

## ALIGNMENTS

### RESULT 1

AG101978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AG101978 717 bp DNA linear GSS 03-NOV-2001  
Pan troglodytes DNA, clone: PTB-105C09.R, genomic survey sequence.

AG101978.1 GI:16722495

GSS.

Pan troglodytes (chimpanzee)

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Unpublished

2 (bases 1 to 717)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Submitted (02-AUG-2001)

Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimpanzee@riken.go.jp, URL: http://hgp.gsc.riken.go.jp/;

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&D process and may have higher chance of

```

clone tracking errors.
PRIMERS
LIBRARY
    Sequencing: M13Rev
    Vector : pKS145
    R.Site 1 : SacI
    R.Site 2 : SacI
    Location/Qualifiers
    1. 717
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /clone="PTB-10509.R"
    /sex="male"
    /cell_type="lymphoblast"
    /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
    Query Match      2.7%; Score 241.6; DB 29; Length 717;
    Best Local Similarity 62.9%; Pred. No. 2e-53;
    Matches 409; Conservative 0; Mismatches 234; Indels 7; Gaps 2;

QY 184 CCGTCTCTGCGGTTTACCGTTTATCTAATTACGGGAACGTGCTGGTGGGAAAGTACC 243
DB 69 CCTTCTGCGCTTTTACGCTATGTATTAACGGGAACGGGGGGTGGCAGAGCACC 128
QY 244 AGCATATCGGCGCTGTACAAATTTAACTGCTGATCAGCGGCGGACCAACATAGCC 303
DB 129 AGCGTTAGTCTGCTCCACACACGATGGACTGCTGCTGCTCAGGGAGCCACAAACGTTGCC 188
QY 304 GCACAGAACCTTCTCCAGACACTCGGAGCTTCTGCTCCACGATCTTCAGCGCTTTGGC 363
DB 199 GCACAGAACCTTCTCCAGACACTCGGAGCTTCTGCTCCACGATCTTCAGCGCTTTGGG 248
QY 364 TTTAAGAGCCGACATCAATATATAGC-----CGTCAGAAAGCTCATCAGACGGAGCC 417
DB 249 TTCAAGAGCCGACATCAATATATAGCAGCGGGGTTCAGCAGTCTGCTGCTTACGGAC 308
QY 418 GTATCCATAGAGCAATTCAGCAACAGAGCTATCGAGTATTCGGCGGTTATAGTGCAC 477
DB 309 GCGGCGCTTAGAGGAACCTCAGCGGGAGCTTGGCCAAATTAATGCTGCGGTTCTCCGAC 368
QY 478 ATTTATGAAGAGGTTTATGGCGAAACCAATGGCATGTACGGGAGCTATATCAACGGC 537
DB 369 ATTGCGCGGAGTTTCAGCGCAACCAAGCCAGGGGGTCTTACTCGGAGTCTCTGGCCCC 428
QY 538 AATTTTGAACCTCTCAGAGATGACGGACCGTGTATGGACTTCCAAATATATGTGA 597
DB 429 GCTTTTGAAGTCTCAGAGACATGACACAGGGGAGCTATGGACCAACCAATGTGATCGTG 488
QY 598 ATCGACGAGCGGAAACCTCTCTTACATATCTACACCGCTGCTGTTTCTTTACTGG 657
DB 489 GTGGACGAGCTGGAACGCTTTCGGTGACATTTCTACAGCGCTGTTTCTTTGCTACTGG 548
QY 658 TTCCTAAACAGTGGCTAAATACCGCTCTTTACCGCCAGGGGGCGGTTCCGTGATAGTA 717
DB 549 TTCTTTCAACGCGTGGCTGCTACCCCACTCTACCGCGCGGCGCGCATTCCTGTCATGTC 608
QY 718 TCGCTCGGTTGCCAACGACAGCAACCGCTTCCAGTCAACTTACAAACCGGAGCGCAA 777
DB 609 TCGTGGGCTCTCCAC-CAGACAGACGCTTTGAGTCTCTCTTTAGCCAGAGACCCAG 667
QY 778 AAGACGAGATATCTCGTGGAGAACATCTTACATTCATGATAGGAA 827
DB 668 GTAAACAAGATTGCGAGTGGGACAAACATCTCCTCCTTCTCTGTTGGGCAA 717

RESULT 2
AG104414
LOCUS
DEFINITION
ACCESSION
VERSION
AG104414.1 GI:16724932

```

```

GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
2
Unpublished
2 (bases 1 to 678)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@sc.riken.go.jp; URL: http://hgp.gsc.riken.go.jp/;
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
    Vector : pKS145
    R.Site 1 : SacI
    R.Site 2 : SacI
    Location/Qualifiers
    1. 678
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /clone="PTB-108F03.R"
    /sex="male"
    /cell_type="lymphoblast"
    /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
    Query Match      2.6%; Score 233; DB 29; Length 678;
    Best Local Similarity 63.1%; Pred. No. 4.3e-51;
    Matches 378; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

QY 184 CCGTCTCTGCGGTTTACCGTTTATCTAATTACGGGAACGTGCTGGTGGGAAAGTACC 243
DB 62 CCTTCTGCGCTTTTACGCTATGTATTAACGGGAACGGGGGGTGGCAGAGCACC 121
QY 244 AGCATATCGGCGCTGTACAAATTTAACTGCTGATCAGCGGCGGACCAACCATAGCC 303
DB 122 AGCGTTAGTCTGCTCCACACACGATGGACTGCTGCTCAGGGAGCCACAAACCGTGGCC 181
QY 304 GCACAGAACCTTCTCGCTCGCTTAAAGAGCTTCTGCTCCACGATCTTCAGCGCTTTTGGC 363
DB 182 GCACAGAACCTTCTCGGAGACTTCCGAGCTTCTGCGCAACCGTCTATAGCGCTTTGGG 241
QY 364 TTTAAGAGCGGACATCAATATAGC-----CGTCAGAAAGCTCATCAGACCGGAGCC 417
DB 242 TTCAAGAGCGGACATCAATATATAGCAGCGGGTTCAGCAGTCTGCTGCTTCTAGGAC 301
QY 418 GTATCCATAGAGCAATTCAGCAACAGAGCTATCGAGTATTCGGCGGTTATAGTGCAC 477
DB 302 GCGGCGCTTAGAGGAACCTCAGCGGGAGCTTGGCCAAATTAATGCTGCGGTTCTCTCCGAC 361
QY 478 ATTTATGAAGAGGTTTATGGCGAAACCAATGGCATGTACGGGACTATATCCAGCGG 537
DB 362 ATTGCGCGGAGTTTCAAGCGCAACCAAGCGGCTTACTCGGAGTCTCTGCGGCC 421
QY 538 AATTTTGAACCTCTCGAGATGACCGGACCGTGTATGAGTCTTCAATATATGTGA 597
DB 422 GCGTTTGAAGTCTTAGAGACATGACACAGGGGAGCTATGACCAACCAATGTGATCGTG 481
QY 598 ATCGACGAGCGGAGAACCGTCTCTTACATATCTACACCGCTGCTGTTTCTTTACTGG 657

```

```

Db      482 GTGACGAAGCTGGACAGCTTTTCGGTGCAATCTTAACAGCCGTGGTCTTTTGCTACTGG 541
Qy      658 TTCTTAACAGCTGGCTAATAATACCCCTCTTTTACCSCCAGGGGCGGTTCCCGTGCAATGTA 717
Db      542 TTCTTCAAGCCCTGGCTGGTACCCACCTCTACCGCGGGGCGGCAATCCCTGCACTGTC 601
Qy      718 TGGTGGCTTGGCAAGCAGACAAACGGTTCCAGTCAACTTACACACGGGACGCA 776
Db      602 TGGTGGGCTCTCCACCCAGACAGACGCTTCTTCACTCTCTCTTTAGCCACGAGACCCA 660

RESULT 3
AG126121
LOCUS      688 bp      DNA      linear      GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-136J09.F, genomic survey sequence.
ACCESSION  AG126121
VERSION     AG126121.1 GI:16655286
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
REFERENCE  1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            BAC end sequences of Library PTB
            Unpublished
            2 (bases 1 to 688)
            Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            Direct Submission
            Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22, Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
            PRIMERS
            Sequencing: -21M13
            LIBRARY
            Vector      : pKS145
            R.Site 1    : SacI
            R.Site 2    : SacI.
            Location/Qualifiers
            1..688
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-136J09.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
Query Match      2.3%; Score 213; DB 29; Length 688;
Best Local Similarity 62.1%; Pred. No. 1.2e-45;
Matches 370; Conservative 0; Mismatches 218; Indels 8; Gaps 2;

Qy      184 CCGTTCCTGGCTTACGGTTATCTAATTACGGGAACTGCTGTGCTGGAAAAGTACC 243
Db      94  CCCCTCTCGCCCTTTTACGCTATGTAAATTACGGGAACGGGGGGCTGGCAAGACACC 153
Qy      244 AGCATATCGGCCCTGTACCAAAATTTAAACTGCCTGATCAAGCGGCGGACCAACCATAGCC 303
Db      154 AGCGTTAGCTGCTCCACCAACACGATGCACTGCCTGTGTCAGGAGCCAAACCGTGGCC 213
Qy      304 GCACAGAACCTATCGCTGGCTTAAGAGCTTCTGTCCACGATCTTCAGCGCTTTGGC 363
Db      214 GCACAGAACCTCTCCACACACTCCGAGCCTACTGCCCAACCGCTATAGCGCTTTGGG 273
Qy      364 TTTAAGAGCCGACATCAATAT-----AGCCGTCAGAAAAGCTCATCAGACCGGAGCC 417

```

```

Db      274 TTCAAGAGCCGCCACATAATATGATGTAGCGGTGAGCAGTCATGTCCTCTACGGAC 333
Qy      418 GTATCCATAGAGCAAAATTACGACACAGAGACTATCGAAGTATTTGGCCGGTTATAGTGGAC 477
Db      334 GCCGCCCTAGAGGAATCTCCAGCGCGGACCTGGCCAAATACTTGGCCGGTACTCTCCGAC 393
Qy      478 ATTATGAAGAGGTTATGGCGAAAAAACCCAAATGGCATGTACGGGACTATATCCAAACGCG 537
Db      394 ATTCCGCCCGAGTTACGGCGAAACCAAGCCCAAGNGGGCTCTACTCGGGAGTCTCTGGCCCC 453
Qy      538 AATTTGAAACCCCTCTCGAAGATGACCGGACCGGTGTTTATGAGACTTCCAATATTTGTA 597
Db      454 GCCTTTGAGTCTCTGAGAGACATGACCAAGNGGCGAGCTATGACCAACCAATGTGATCGTG 513
Qy      598 ATCGACGAGCGCGAACCCTGCTCTCTTACATATCTACCAACCGTCGCTGTTCTTTTACTGG 657
Db      514 GTGACCAAGCTGGAACGCTTTTCGTGCAATNCTAACAGCCGGTGTCTTTTGTACTACTGG 573
Qy      658 TTCTTAACAGCTGGCTGGTACCCCTCTTTTACCGCCAGGGGGGGGTTCGGTGATAGTA 717
Db      574 TTCTTTCAAGCGCTGGCTGGTACCC--ACTTACCGCGGGCGGCATCCCTGTCATCGTC 631
Qy      718 TGGTGGCTGGCGCAACGACGACAAACGCTTCAGTCAACTTACAACCCAGCGAC 773
Db      632 TGGTGGGCTCTCCACCCAGACAGCGCTTTTCAGTCTCTCTTTAGCCACGAGAC 687

RESULT 4
AG051826/c
LOCUS      721 bp      DNA      linear      GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-033P12.R, genomic survey sequence.
ACCESSION  AG051826
VERSION     AG051826.1 GI:16589269
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
REFERENCE  1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            BAC end sequences of Library PTB
            Unpublished
            2 (bases 1 to 721)
            Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            Direct Submission
            Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22, Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
            PRIMERS
            Sequencing: M13Rev
            LIBRARY
            Vector      : pKS145
            R.Site 1    : SacI
            R.Site 2    : SacI.
            Location/Qualifiers
            1..721
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-033P12.R"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source

```

```

RESULT 5
AGI24918
LOCUS      695 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-135A07.R, genomic survey sequence.
ACCESSION AGI24918
VERSION    AGI24918.1 GI:16654083
KEYWORDS   GSS
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 685)
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suhei-ro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT    Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY    Vector : pKS145

```

Unpublished  
2 (bases 1 to 656)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1. .656  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-033P12.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
ORIGIN  
Query Match 1.2%; Score 111; DB 29; Length 656;  
Best Local Similarity 53.5%; Pred. No. 5.7e-18;  
Matches 300; Conservative 0; Mismatches 230; Indels 31; Gaps 2;  
QY 955 GAAGTCATGAGTTACGTPAGACAGATTGTCGTTCTCTAAAGTAAGATTATGGATCCTTTA 1014  
Db 90 GACATCCTGCGCTACGTCGACCGCTTCGTCGTCGCCAGCGGCCCAATATGACCCCGCG 149  
QY 1015 GAATACATTGGCTCGACCGCACTCTTTTGTACATAGCAGGTAAGGGGTATCTAA 1074  
Db 150 CAGTACGTGGCTGGACCGCGCTCTTCCTCTCCACGCCGAGTCAAGACTTCTCCAC 209  
QY 1075 AACCTACACATGTCCTAAAGCTA-----GGGGG 1104  
Db 210 ACCCTCATGCGCAGCTCAAGACTCGCGGCGGGGGTCCCGGGGGAACCGGGGA 269  
QY 1105 GATACAGGACACGAAGCTTTTACCTGTCCTGCGTGTGCGAGGTGTTTGAAGCCG 1164  
Db 270 GATGGGGCGGGGTGACCATGTTTACCTGCGCGTGGAGTGCAGAGTCTTCTCGATCCC 329  
QY 1165 TTTGAGGATACAGCGGGCGCTCACTCACCACCTCACCCTGACCGAGTGGGTGACA 1224  
Db 330 CTGGCCAGTACAGACCTGGTTCGGCTTCGGGCTTCGGGCTTCAGGCCACACCTGGCTCCA 389  
QY 1225 AAAAACCTCTTTAAGTTAAGTAATTATTCGAGTTTGTGGACCGAGGACATGTCATGTC 1284  
Db 390 AAGAAATTATGACGCTCGGTAACTACTCGCAGTTTCGCCGACCAAGACATGTTCCGGGTG 449  
QY 1285 GCGACGGA-ATCCTGACGTTTCGACCCAGTTTACCTTTTATCACCAGTTTGTATAAAA 1343  
Db 450 GGCACCGAGGAGGACGAAGAACCGGTGAAAGGTCACTTACATGTCACCTACGTCAGCA 509  
QY 1344 CAGCCACGTGTCTTTAAATGAAAAAACAATAATGATATGTGGTTCAGGGTACGTA 1403  
Db 510 CAGCTCGGTGTGCTGCTCACTCAAGACCAAAAGTCCATCTGTGGCTATACCGGCACCT 569  
QY 1404 CTTTGAGTTTAAAAGATTCCTAGACAGGAACTCTTGTGGAGACGATTCGCGAGACCG 1463  
Db 570 CGGGGACTTTTGGACACGCTCGAGGGGACAGCTTCGCGAGGCCCAACCGGATGAGCA 629  
QY 1464 TCCGGAATACGTATACGGTTT 1484  
Db 630 GTCCGAGTATGTCTACAGCTT 650

RESULT 7  
AJ395572/C  
LOCUS AJ395572.1 696 bp mRNA linear EST 25-JAN-2001  
DEFINITION AJ395572 dkf426 Gallus gallus cDNA clone 22p14r1, mRNA sequence.  
ACCESSION AJ395572  
VERSION AJ395572.1 GI:7126625  
KEYWORDS EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 696)  
Abdrakhmanov, I., Lodygin, D., Geroth, P., Arakawa, H., Law, A.,  
Plachy, J., Korn, B. and Buerstedde, J.M.  
A large database of chicken bursal ESTs as a resource for the  
analysis of vertebrate gene function  
Genome Res. 10 (12), 2062-2069 (2000)  
20568495  
PUBMED 11116100  
COMMENT Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.  
Location/Qualifiers  
1. .696  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="CB"  
/db\_xref="taxon:9031"  
/clone="22p14r1"  
/tissue\_type="Bursa of Fabricius"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_lib="dkf426"

ORIGIN  
Query Match 1.2%; Score 105.8; DB 9; Length 696;  
Best Local Similarity 51.4%; Pred. No. 1.5e-16;  
Matches 270; Conservative 0; Mismatches 25; Indels 3; Gaps 1;  
QY 3634 AGCCAGGCGAGACGATGGCGCGCTTGAGGACGAGATGTCTCTCCGTTTATCAGCGT 3693  
Db 601 AGAGAAAGGGAGGGGTGAACCGCTCGAGACGCTGGTGGCTTCTCTGTCATAGAGCT 542  
QY 3694 CGCCTTGTCTATAGCTTTCTTCCCGACAGATATAAAGCGACCTCTGAAGCTTGGCAGA 3753  
Db 541 GCGCTTCTTCTGTGATAGCTCCCGACAGATGAAACCAACCCCATGCAGGTTCTTGT 482  
QY 3754 AAGGCGAGCTTATATGTAGTGGTGAACCAATCCAGCAAGGTTGGCGTGTGACCCCGC 3813  
Db 481 GAGCCAGGAGACGACCATCCGTGAATCTGCCAGCCCTCTCTTGTGGGACGTGGC 422  
QY 3814 CTTCCCTCTTTCACCGTCAGAAATGTTTAAAGAGCAACTCCCGCTTTTGGCCACA 3873  
Db 421 CTGTGGGCTCGACCGTGGAGACGCGCTTGAGCAGGAGCACCCCTCTTGGCCCGC 362  
QY 3874 GTCTAAGCACCGCTGAGAGGCGCTGAACCCAGGTACAGTATTTGGCTATCTCTTTAAA 3933  
Db 361 AOTCAGATCGCGCTGACCGAGATGGTGAAGTCTTCATATACGTCGACAGTCTTTGTA 302  
QY 3934 AATATTTTGGAGCTTGGAGGTATATCTGTAGTCCGAGCAACGCTAAAGGCTAGTCCGGT 3993  
Db 301 AATGTTTCCAACTGGGGGAGGTGGAACACAGGTTTCTGGACACTGAAACAGAGCCCATG 242  
QY 3994 GCGCTG---ACCGCATGGTAGGATCTGGCCAGGATCAGACTTTTAAATCTCTCCGG 4050  
Db 241 AGCTTGATTAGGTCCATGATAGGATCTTGTCCCGAATTCACTTTCACATCCCTGAT 182  
QY 4051 TTCCGCAACATACGACCCACCATAAATCTATCTATCGGTGGATAAATCAGGTTACGTT 4110

```

Db      181  GTCCACATCTGTGTCAGTGAAGCTCTCGGGGGGGGTACACCGTGTACCTCTT 122
Qy      4111  GTCATATCATTAACGGCGCTTACGACGAGCGCGGAGCTTCTGTTT 4155
Db      121  CCTCTCCTCGGCACGAATGCCATCAGCTCCACGAAGTAGGGCTT 77

RESULT 8
LOCUS   BU444447/c
DEFINITION 603213237f1 CSOREN13 Gallus gallus cDNA clone CHEST195j9 5', mRNA
sequence.
ACCESSION BU444447
VERSION   BU444447.1 GI:25933758
KEYWORDS EST
SOURCE   Gallus gallus (chicken)
ORGANISM Gallus gallus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
          Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 911)
          Boardman,P.E., Sans-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
          Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
          A Comprehensive Collection of Chicken cDNAs
          Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
CONTACT: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

```

## FEATURES

```

Source
1..911
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="CHEST195j9"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN13"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research
(1996) 7: 791, except that a significantly longer
reannealing hybridization was used."

```

## ORIGIN

```

Query Match
Best Local Similarity 51.1%; Score 103.6; DB 13; Length 911;
Matches 268; Conservative 0; Mismatches 254; Indels 2; Gaps 1;

Qy      3634  AGCCAGGGGAGACGATCGCGCGCTTGAGACGAGATGTTCTGTCGGTTATCAGCGT 3693
Db      693   AGAGAAAGGGGAGGGTGAACCGCTGCGAGGACGTGTTGCGGTTTCTCTGTCATAGAGCT 634
Qy      3694  CGCCTGTCTATAGCTTCTTCCACAGCATAAACGACCTGTAAGCTTGGCAGA 3753

```

```

Db      633   GCCCTTCTTGTGTATAAGCTCCCCACAGCATGAAACACCCCATGACGGTTCTTGT 574
Qy      3754   AAGCAGCTTATTATGTAGTGGTGAACCAATCCAGCAAGGTTGGCGTGTGACCCGCG 3813
Db      573   CAGCAGGAGAGACCAACATCCGTGAAGTGTCTCCAGCCCTTCTCCTTGTGGAGTGGC 514
Qy      3814   CTTCCCTCTTTCCACCGTCAGAAATGGTGTAAAGAGAGAACTCCCGTTTGGCCCAACA 3873
Db      513   CTGGTGGGCTCGCACCCGTCAGGACGGGCTTGAGCAGGAGCACACCGCTGCTTGGCCAGCC 454
Qy      3874   GTCATAGCACCCGTCAGAGAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAA 3933
Db      453   AGTCAGATCGCCGTCGACCCAGGATGGGTGAAGTCTTCAATATCAGTCGACACTTTTGTGA 394
Qy      3934   AATATTTTGA--GGCTTGGAGGTATAGTGTAGTCCGAGCAACGCTAAAGGCTAGTCCG 3991
Db      393   AATGTTTTCCAACTGGGGGAGGTGGAACAGGTTTCTGACACTGAACACAGAGCCCATGA 334
Qy      3992   GTGGCTTGACCGCGATGTTAGGATCTCTGGCCAGGATCAGACTTTAATATCTCCCGGT 4051
Db      333   GCTTGATTAGGTCCATGATTAAGGATCTTGTCCAGAAATTACAATTTTACATCCCTGATG 274
Qy      4052   TCGCAACAAATACGACCAACACATAATTCTATCTATCGGTGGATAAATCACGGTTACGTTG 4111
Db      273   TCGCACATCTGTGTCGAAGTGAAGACCTGCTCCGGGGCGGGTACACCGTGTACCTCTTC 214
Qy      4112   CTCATATCCATAACGGCTTTCAGCAGAGCCGCGAGCTTCTGTTT 4155
Db      213   CTCCTCTCGGCACACGATGCCATCAGTCCACGAAAGTAGGGCTT 170

```

## RESULT 9

```

LOCUS   BX706705/c
DEFINITION BX706705 XGC-tadpole Silurana tropicalis cDNA clone TTPA010a21 5',
          mRNA sequence.
ACCESSION BX706705
VERSION   BX706705.1 GI:38368912
KEYWORDS EST.
SOURCE   Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
          Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 987)
          Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
          Sanger Xenopus tropicalis EST project 2001 (11_2003)
          Unpublished (2003)
          Contact: Croning MDR
          Sanger Institute
          Hinxton, Cambridgeshire, CB10 1SA, UK
          Email: trop@sanger.ac.uk
          TROPICALIS_SEQUENCE ID: TTPA010a21.pikSP6
          Sequencing primer: SP6
          This sequence is from a Xenopus Gene Collection (XGC) library
          constructed by Nigel Garrett.
          cDNA was oligo dt primed from sug of poly A+ RNA from tadpole
          embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
          EcoRI at the 5' end and NotI at the 3' end.
          Vector: pCS107; Site 1: EcoRI; Site 2: NotI
          Host: Escherichia coli DH10B.
          Location/Qualifiers
            1..987
              /organism="Silurana tropicalis"
              /mol_type="mRNA"
              /db_xref="taxon:8364"
              /clone="TTPA010a21"
              /dev_stage="tadpole (stage 35-40)"
              /lab_host="E. coli DH10B"
              /clone_lib="XGC-tadpole"
              /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
              was oligo dt primed from sug of poly A+ RNA from tadpole

```

## FEATURES

Source



embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 1.1%; Score 103.6; DB 13; Length 987;  
 Best Local Similarity 50.9%; Pred. No. 7.1e-16; Indels 3; Gaps 1;  
 Matches 271; Conservative 0; Mismatches 259; Indels 3; Gaps 1;  
 QY 3631 GGTAGCGAGGAGCGGATGCGCGCTTGAGGACGAGATGCTCTGCGGTTATACAG 3690  
 DB 881 GGCTGAGAGGAGAGGGTGCACAGTTTGGAGAACAGGTGCTCTTCTGTCAATATT 822  
 QY 3691 CGTGGCTTGTCTATAGCCTTTTCCCCACACATATAAAGCGACCTTGAGCTTGGC 3750  
 DB 821 GCTGCTTTCTCTGAGCATTAAGTCCCAACATGAAACCAACCAACCATGATGTTT 762  
 QY 3751 AGAAGGCGCTTATTATGATGCTGTGTAACCAATCCAGCCAAAGGTGGCGGTGACCC 3810  
 DB 761 GTTAGCCAGGACACACAGAGTCTGTGAATCTTCCATCCACAGTCTTGTGAGATT 702  
 QY 3811 CGGCTTCCCTTTTCCACCTGAGATGAGTGTGTTTAAAGCAGAACTCCCGTTTCCCA 3870  
 DB 701 TGCATTATGAGCTCGGACTGTGATCAGCATTTGAGAGGAGGACTCTTGTCTGGCCCA 642  
 QY 3871 ACAGTCTAAGCACCCTGAGAGGCGGTGAACCCAGGTACAGTATTGGCTATCTTT 3930  
 DB 641 GCCTGTGAGATCACATGCGCTGGTGTGAATCTTCTCATATCAGTTTCCAGTTCTTT 582  
 QY 3931 AAAAATATTTTGGAGCTTGGAGTACTGTAGTCCGAGCAACGCTAAAGCTAGTCC 3990  
 DB 581 GTACATATTCACTAAGCTTGGAGGAGGAGGAGTGTCTTAACGCTAAAGCAAGTCC 522  
 QY 3991 GGTGGCTTGA---CCGGAGGTGAAGATCTCTGGCAAGATCAGCACTTATATATCTTC 4047  
 DB 521 ATGAGCTTGTATTTGGTCCATGATAGGATCTTGGCCCAAAATGACAACTTCACTTT 462  
 QY 4048 CGGTTCGCAACAATPACACCAACACATATTTCTATCTATCGTGGATATAATCAGTTAC 4107  
 DB 461 GATGTCACCAATTTGTCTCCAGGTGAATACCTCTTCTGGGGTGGATAAATCTGTATT 402  
 QY 4108 GTTCTCATATCCATACCGCTTACGAGCGCGGAGCTTCTGTTTAAATA 4161  
 DB 401 CTTTCTCTCTCTCCAAATAATTCGATAGCTTAAACAAAGTAGGGCTTAGCAAA 348

RESULT 10  
 BX719507/c  
 LOCUS BX719507 XGC-tadpole silurana tropicalis cDNA clone TTPA030n15 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BX719507  
 VERSION BX719507.1 GI:38392248  
 KEYWORDS EST.  
 SOURCE Silurana tropicalis (western clawed frog)  
 ORGANISM Silurana tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Silurana.  
 1 (bases 1 to 871)  
 Crounig, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.  
 Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 Unpublished (2003)  
 Contact: Crounig MDR  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE ID: TTPA030n15.plkSP6  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Nigel Garrett.  
 cDNA was oligo dT primed from 5' of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end.  
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
 Host: Escherichia coli DH10B.  
 Location/Qualifiers

FEATURES

source  
 1..871  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clones="TTPA030n15"  
 /dev\_stage="tadpole (stage 35-40)"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="XGC-tadpole"  
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dT primed from 5' of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107  
 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 1.1%; Score 103.2; DB 13; Length 871;  
 Best Local Similarity 50.7%; Pred. No. 8.6e-16;  
 Matches 271; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
 QY 3631 GGTAGCGAGGAGCGGATGCGCGCTTGAGGACGAGATGCTCTGCGGTTATACAG 3690  
 DB 864 GGCTGAGAGGAGAGGGTGCACAGTTTGGAGAACAGGTGCTCTTCTGTCAATATT 805  
 QY 3691 CGTGGCTTGTCTATAGCCTTTTCCCCACAGATATAAAGCGACCTCTGAAGCTTGGC 3750  
 DB 804 GCTGCTTTCTCTGAGCATTAAGTCTCCCAACATGAACCAACCAATCATGTTTTT 745  
 QY 3751 AGAAGGCGAGCTTATTATGATGCTGTGTAACCAATCCAGCGCAAGGTGGCGGTGACCC 3810  
 DB 744 GTTAGCCAGGACACCAAGAGTCTGTGAATCTTCCATCCACAGTCTTGTGAGATT 685  
 QY 3811 CGCTTCTCTCTTCCACCGTCAGATGCTGTGTTAAAGCAGAACTCCCGTTTGGCCCA 3870  
 DB 684 TGCATTATGAGCTCGATCTGTGATCAGCATTTAGGAGGAGGACTCTTGTGCGCCCA 625  
 QY 3871 ACAGTCTAAGCACCGCTGAGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTT 3930  
 DB 624 GCCTGTGATGATGCTTCCATGATAGGATCTTACCCAAATGACAACTTCACTTTT 565  
 QY 3931 AAAAATATTTTGGAGCTTGGAGGTACTGTAGTCCGAGCAACGCTAAAGGCTAGTCC 3990  
 DB 564 GTACATATTCTAAGCTTTGGAGGAGGGGAGTCTTCTTAAGCTTAAGCAAGTCC 505  
 QY 3991 GGTGGCTTGA---CCGGAGGTGAAGATCCTGCGCAAGGATCAGACTTTTAATTCCTC 4047  
 DB 504 ATGCTGTTGATTTGGTCCATGATAGGATCTTACCCAAATGACAACTTCACTTTT 445  
 QY 4048 CGGTTCGCAACAATPACACCAACACATATTTCTATCTATCGTGGATATAATCAGGTTAC 4107  
 DB 444 GATGTCACCAATTTCTGTCCAGGTGAATACCTTTCTGGGGTGGATATAAATCTGTATT 385  
 QY 4108 GTTGTCTCATATCCATAAGCGCTTCCAGAGAGCGCGGAGCTTCTGTTTAAAAA 4161  
 DB 384 CTTTCTCTCTCTCCAAATAATTCGATAGCTTAAACAAAGTAGGGCTTAGCAAA 331

RESULT 11

AL894178/c  
 LOCUS AL894178 XGC-egg silurana tropicalis cDNA clone Tegg092j11 5', mRNA  
 DEFINITION sequence.  
 ACCESSION AL894178  
 VERSION AL894178.2 GI:38694463  
 KEYWORDS EST.  
 SOURCE Silurana tropicalis (western clawed frog)  
 ORGANISM Silurana tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Silurana.  
 1 (bases 1 to 645)

AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.  
TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
JOURNAL Unpublished (2003)  
COMMENT On Sep 16, 2002 this sequence version replaced gi:22944729.  
Contact: Taylor R  
Sanger Institute  
Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: tropesanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS\_SEQUENCE\_ID: TEGG092j11.plkSP6  
Sequencing primer: SP6  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn.  
cDNA was oligo dT primed from Sug of poly A+ RNA from egg.  
ECORI-NotI cut cDNA was then ligated into pCS107 with EcorI at the  
5' end and NotI at the 3' end.  
Vector: pCS107; Site 1: EcorI; Site 2: NotI  
Host: Escherichia coli XL1-blue.  
Location/Qualifiers  
1..645  
/organism="Silurana tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="TEGG092j11"  
/dev\_stage="egg"  
/lab\_host="Escherichia coli XL1-blue"  
/clone\_lib="XGC-egg"  
/note="Vector: pCS107; Site 1: EcorI; Site 2: NotI; cDNA  
was oligo dT primed from Sug of poly A+ RNA from egg.  
ECORI-NotI cut cDNA was then ligated into pCS107 with  
EcorI at the 5' end and NotI at the 3' end"

FEATURES  
source  
1..645  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:7012130"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHD XGC Kid1"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.2 kb. Constructed by Life  
Technologies. Note: This is a Xenopus Gene Collection  
(XGC) library."

ORIGIN  
Query Match 1.1%; Score 102.4; DB 14; Length 699;  
Best Local Similarity 50.7%; Pred. No. 1.3e-15;  
Matches 269; Conservative 0; Mismatches 255; Indels 3; Gaps 1;  
QY 3638 AGGGAGACGATGGCGCGCTTGAGACGAGATGCTCTGCGTTTATCAGCGTCGCC 3697  
Db 640 AGGGAGAGGGGTGCACAGTTTGAGAACAAAGTGCCTCTTCTGTCATATTGCTGCCT 581  
QY 3698 TTGCTATAGCTTCTTCCCCACAGATATAAAGCGACTCTGAAGCTTGGCAGAAAGG 3757  
Db 580 TTCTCTGAGATAGCTCCCAACATGAAACCAACATGATGATGTTTGTGTTAGC 521  
QY 3758 CAGCTTATATGATGTTGAACCAATCCACGCAAGGTGGCGGTGACCCCGCTTC 3917  
Db 520 CAGGACACCAAGAGTCTGTGAAGTCTCCATCCACAGTCTTGTGAGAGTTTGCATTA 461  
QY 3818 CCTCTTCCACGTCAGATGGTCTTAAAGCAGACTCCCGTTTGGCCACAGTCT 3877  
Db 460 TGAGCTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 401  
QY 3878 AAGCACCCTGAGAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAAAATA 3937  
Db 400 AGATCACCATGCCCTGGGTGTGAAATCTTCAATATCAGTTTCCAGTTCTTTGTACATA 341  
QY 3938 TTTTGTAGGCTTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3997  
Db 340 TTCACTAAGCTTGAGGAGGGGGGACTGGTTTCTTAACGCTAAAGCAAGTCCATGTGCT 281  
QY 3998 TGA---CGCGATGTTAGGATCTTGCGCAAGGATCAGCACTTTAATATCTCCGGTTCG 4054  
Db 280 TGATTTGTGTCATGATAGGATCTTGACCCAAATGACACCTTCCATCTTTGATGTCA 221  
QY 4055 CAACAATACGACCAACATATTTCTATCTATCGGTGATTAATACAGGTGATGCTGCTC 4114  
Db 220 ACCATTTGTGCCAGTGAATACCTTCTCTGGGGGTGATAAACTGTATTTCTTTCTC 161  
QY 4115 ATATCCATAAGCGCTTCAGCAGAGCGCGAGCTTCTGTTTAAAAA 4161  
Db 160 TCTTCTGCACAAATTCATAGCTTAACAAGTAGGCTTAGCAAA 114

RESULT 12  
CF521628/c  
LOCUS CF521628.1 699 bp mRNA linear EST 10-SEP-2003  
DEFINITION AGENCOURT\_15508682 NICHD\_XGC\_Kid1 Xenopus laevis cDNA clone  
IMAGE:7012130 5', mRNA sequence.  
ACCESSION CF521628  
VERSION CF521628.1 GI:34572501  
KEYWORDS EST.  
SOURCE Xenopus laevis (African clawed frog)  
ORGANISM Xenopus laevis  
REFERENCE 1 (bases 1 to 699)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Igor Dawid  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL4721 row: m column: 24  
High quality sequence stop: 634.  
Location/Qualifiers  
1..699  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:7012130"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHD XGC Kid1"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.2 kb. Constructed by Life  
Technologies. Note: This is a Xenopus Gene Collection  
(XGC) library."

ORIGIN  
Query Match 1.1%; Score 102.4; DB 14; Length 699;  
Best Local Similarity 50.7%; Pred. No. 1.3e-15;  
Matches 272; Conservative 0; Mismatches 261; Indels 3; Gaps 1;  
QY 3638 AGGGAGACGATGGCGCGCTTGAGACGAGATGCTCTGCGTTTATCAGCGTCGCC 3697  
Db 613 AGAGGAGAGGGTGCACGCTTGAAGAACATGGTGCCTTTTCTGTCATATTGCTGCT 554  
QY 3698 TTGCTATAGCTTCTTCCCCACAGATATAAAGCGACTCTGAAGCTTGGCAGAAAGG 3757  
Db 553 TTCTCTGAGATAGCTTCCCAACATGAAACCAACCATCCATGCTTCTTATTAGC 494  
QY 3758 CAGCTTATATGATGCTGTAACCAATCCACGCAAGGTGGCGGTGACCCCGCTTC 3817  
Db 493 CAGGATACCAACATCTGTGAAGTGTTCCTCCATCCACAGTCTTGTGAGAGTTGCATTA 434  
QY 3818 CCTCTTTCACCGTCAGATGGTGTGTTAAAGCAGAACTCCCGTTTGGCCAAAGTCT 3877  
Db 433 TGAGCTCGACTGTGACACAGATGAGGAGGAGTCTCTTGTGGCCAGCTGTG 374  
QY 3878 AAGCACCCTGAGAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAAAATA 3937  
Db 373 AGATCCCATGACTCTGGCGTGAATAATCTTCGATATCGGTTTCTAGTTCTTTGTACATA 314  
QY 3938 TTTTGTAGGCTTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3997

313	TTCACTAAGCTGGGTGGAGTGGCTTCTTGACACTAAAGCAAGACCACGAGCT	254
3998	TGA---CCGCGATGGTAAAGATCCTGCCAAGATCAGACCTTTAATATCTCCGGTTGCG	4054
253	TGATTTGGTCCATGATAAGGATCTTGGCCTAAAAATAACAACCTTTACATCTTTTGATGTCA	194
4055	CAACAATACGACCAACACATAATTCTATCTATCGGTGGATAAATACGGTTACGTTGCCTC	4114
193	ACCATTTGTCTCCAGGTGAATAGCTCTCTGTGGTGGATAAACTGTGCATTTCTTTCTC	134
4115	ATATCCATACCGCGCTTACGACAGCGCGAGCTTCTGTTTTAAAAATGGAGAAAG	4170
133	TCCTTCGCATGAAATTTGTAGCTTGACAAAGTAGGGTTTTGCAAACTCAGCAAG	78

RESULT 13	LOCUS	AL669156	806 bp	mRNA	linear	EST 14-JAN-2002
AL669156/c	DEFINITION	AL669156	directional larval cDNA library	Ciona intestinalis	cDNA	
			clone 043ZC09 5', mRNA sequence.			

ACCESION	AL669156
VERSION	AL669156.1
KEYWORDS	GI:18142413
SOURCE	EST.
ORGANISM	Ciona intestinalis
	Ciona intestinalis
	Ciona intestinalis
	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
	Phlebobranchia; Cionidae; Ciona.
REFERENCE	1 (bases 1 to 806)
AUTHORS	Genoscope.
TITLE	Ciona intestinalis directional larval cDNA library
JOURNAL	Unpublished (2002)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 19 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr IMPORTANT: this sequence may contain errors. The Ciona intestinalis library from which the clone was isolated may be contaminated with cDNAs from bacteria or other Eukarya. Directional larval cDNA library originate from Dr. M.Branno, Stazione A.Dohrn, Naples, Italy, and was prepared in pBluescript2skt+.

```

FEATURES
    source
        Location/Qualifiers
            1..806
                /organism="Ciona intestinalis"
                /mol_type="mRNA"
                /db_xref="taxon:7719"
                /cclone="0432C09"
                /clone_lib="directional larval cDNA library"
                /note="vector: pBluescript2SK+"

```

ORIGIN	Query Match	1.1%;	Score 101;	DB 9;	Length 806;
	Best Local similarity	54.5%;	Pred. No. 3.3e-15;		
	Matches 24;	Conservative 0;	Mismatches 198;	Indels 4;	Gaps 2;
Qy	3658	CTTGAGACGAGATGTCTCTCTCCGTTTATCAGCGTCGCCCTTGCTATAGCCTTTCCTCC	3717		
Db	729	CTTCAGACACAGGTGACGTTCTCTGGTCTATTATTTTCTCCCTCTCTGTGAGATAACTNCC	670		
Qy	3718	CCACAGCATAAAC-GCAGCTCTGAAGCTTTGGCAGAAAGCAGCTTATTATGTAGCTGG	3776		
Db	669	CCACAGCNATAACACCAACCCCGTTTAACTTGGTGTCTATCCATGAATGACGGAATCTG	610		
Qy	3777	TGAACCAATCCAGCCAAAGTTGGGCTGTGACCCCGCCTTCCCTCTTTCCACCGTCAGAA	3836		
Db	609	TCAACTTCTCCCATCCCTTATCTTATGTAGTTAGGTTCCACCATCTCTTACTGTTAACA	550		
Qy	3837	TGGTGTTTTAAAGCAGAACTCCCGTTTGGCCAAACAGCTTAAGCACCCCGTGAGAAGGAG	3896		
Db	549	CGGCATTGAGAGCAGAACACCCCTGATCAGCCCAACCCCTGTTTAAATCTCCATGGTCTGGG	490		
Qy	3897	CGGTGAACCCAGGTACAGTATTGGCTATCTCTTTTAAAAATATTTTGGGGTTGGAGGTA	3956		

489	TATTGAATCCTTCAATATACCTTTCAGTCTTTGAAACATGTTNTTAAGCTGGTGGAG	430
Db		
3957	TACTGTAGTCCGGAGCAACGCTAAAGGCTAGTCCGGTGGCTTGACC---GCCATGGTTAAG	4013
Qy		
429	TACGTATGCCCTTTCTGAAACACTAAACATAAACCATGAGCTGTCTCAGGCCCATGGTAAG	370
Db		
4014	GATCCTGGCCAGGATCAGACCTTTAATATCTTCGGTTCGCAACAATACGACCACCACA	4073
Qy		
369	GGTCTTGCCCAAGTATTACACTTTGATGCTTCAATTTTACAATGTGTGTCCATGCAA	310
Db		
4074	TAATTCCTATCTATCGGTGGATAAA	4097
Qy		
309	ACACCTGGTGTCTGGAGGGTAAA	286
Db		

RESULT 14					
BI940552/c	BI940552	574 bp	mRNA	linear	EST 18-OCT-2001
LOCUS	db71f06.y1	Wellcome CRC psk egg xenopus laevis cDNA clone			
DEFINITION	IMAGE:3378467	5' similar to SW:UNG_HUMAN P13051 URACIL-DNA GLYCOSYLASE PRECURSOR ; mRNA sequence.			

ACCESSION	BI940552
VERSION	BI940552.1
KEYWORDS	GI:16255024
SOURCE	EST.
ORGANISM	Xenopus laevis (African clawed frog)
	Xenopus laevis
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.

REFERENCE	1 (bases 1 to 574)
AUTHORS	Clifton, S. Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
TITLE	WashU Xenopus EST project, 1999
JOURNAL	Unpublished (1999)
COMMENT	Other ESTs: db71f06.x1 Contact: Sandy Clifton, Ph.D. WashU Xenopus EST project, 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute). DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/[L1NL](http://L1NL) at: [info@image.l1nl.gov](mailto:info@image.l1nl.gov)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 454.

FEATURES	source
Location/Qualifiers	
1..574	
/organism="Xenopus laevis"	
/mol_type="mRNA"	
/db_xref="taxon:8355"	
/clone="IMAGE:3378467"	
/tissue_type="egg"	
/lab_host="DH10B (phage-resistant)"	
/clone_lib="Wellcome CRC psk egg"	
/note=vector: plusscript SK-; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Library was constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute). Note: This is a Xenopus Gene Collection (XGC) library."	

ORIGIN	Query Match	1.1%	Score 100.8	DB 12	Length 574
	Best Local Similarity	50.6%	Pred. No. 3.2e-15		
	Matches 271; Conservative	0	Mismatches 262	Indels 3	Gaps 1
Qy	3638	AGGGGAGACGGATGCGCGCGCTTGAGGACGAGATGTCCTCTCTCGCTTTATCAGGCTCCGC	369		

558 AGAGGAGGGTGCACGGTTTGAGAACATGCTGCTTTTCTGTCATATTGCTGCT 499  
3698 TTGCTATAGCTTTCTTCCGCCAGCATATAAAGCAGCACTCTGAAGCTTGGCAGAAAG 3757  
498 TTCTTCTGGCAGTAGGCTCCGCCAGCATGAAACCAACCCATCCATGTTCTTATTAGC 439  
3758 CAGCTTATTATAGCTGGTGAACCAATCCAGCAAGGTTGGCGTGACCCCGCTTC 3817  
438 CAGATACCAACATCTGTGAACTGTTCCATCCACAGTCTCTGTGAGAGTTGCAATTA 379  
3818 CTTCTTTCCACCGTCAGAAATGGTGTGTTAAAGCAGCACTCCCGTTTGGCCCAACAGTCT 3877  
378 TGAGCTCGCACTGTCCAGCAGCATTGAGGAGGAGCTCTCTGTTTGGCCAGCCTGTG 319  
3878 AAGCACCCGTCAGAGAGCGGTGAACCCAGGTACAGTATGCTATCTCTTAAATAA 3937  
318 AGATCCCATGACCTGGCGGTGAATCTCTCGATATCGGTTTCTAGTTCTTTGTACATA 259  
3938 TTTTGTAGGCTTGAGGTATAGTGTAGTCCGAGCAACGCTAAAGGCTAGTCGGGTGCT 3997  
258 TTCACTAAGCTGGGTGGAGTGGGACCTGCTTCTTGACACTAAAGCAAGACCATGAGCT 199  
3998 TGA---CCGGATGATGATGATCTGCGCAGGATCAGACTTATATCTCCCGTTCG 4054  
198 TGATTTGGTCCATGATAGGATCTGGCCCTAAATAACACCTTACATCTTTGAGTCA 139  
4055 CAACAATACGACCAACCATATAATCTATCTATCGGTGATAAATCAAGTTCAGTTGCTC 4114  
138 ACCATTGTTGTCAGAGTGAATACCTCTCTGTTGGTGAATAAAGTGTGCAATTTCTTC 79  
4115 ATATCCATAGCGCTTCAGAGAGCGCGAGCTCTGTTTAAATGAGAGAAAG 4170  
78 TCTTCTGCAATGAAATTTGATAGCTTGACAAAGTAGGGTTTGCAAACTCAGCCAG 23

RESULT 15  
CD321504/c  
LOCUS  
DEFINITION  
StrP538.002866 Sea urchin embryo 7hr cleavage stage cDNA library  
MPGp538 Strongylocentrotus purpuratus purpuratus cDNA clone  
CALTP538D1025;MPI\_538\_25D10 5', mRNA sequence.  
CD321504  
CD321504.1 GI:34793565  
EST.  
Strongylocentrotus purpuratus  
Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoida; Euechinoida; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
1 (bases 1 to 687)  
Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,  
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.  
Generation, annotation, evolutionary analysis, and database  
integration of 20,000 unique sea urchin EST clusters  
Genome Res. 13 (12), 2736-2746 (2003)  
Contact: Poustka AJ  
Laboraty 145, Dept Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: poustka@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting  
(ONF) to reduce sequencing redundancy. According to the ONF  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 smer oligonucleotides are grouped into clusters. One  
clone per ONF cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well as the coordinates of  
the other clones assigned to the same ONF cluster as the clone from  
which the above EST is generated is available at the sea urchin  
project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/) . cDNA

clones and filters are distributed via the Resource Center/Primary  
Database of the German Human Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGAGCTTTACACTTATCTCTCCGGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGACAGCTGGGAGAGGGATGTG 3' (M13FSP) 3'-seq  
Seq primer: 5'-CGGTCCGGAATTCGGGT-3', pSP03/86  
High quality sequence stop: 687.  
Location/Qualifiers  
1. 687  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="CALTP538D1025;MPI\_538\_25D10"  
/tissue type="whole embryo"  
/dev stage="embryonic 7hr"  
/lab host="E. coli, X11 blue"  
/clone\_lib="Sea urchin embryo 7hr cleavage stage cDNA  
library MPGp538"  
/note="vector: pSport1; Site 1: NotI; Site 2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-pCATAGTTCATGATCGGAGCGGCC (T)15-3' and a  
SalI 5'-TCGACCCACGGCTCG-3' adapters (Gibco BRL)"

ORIGIN

Query Match 1.1%; Score 100.8; DB 14; Length 687;  
Best Local Similarity 52.7%; Pred. No. 3.4e-15; Indels 3; Gaps 1;  
Matches 235; Conservative 0; Mismatches 208;  
QY 3656 GCCTTGAAGGAGAGATGCTCTGTCGGTTTATCAGCGTCGCTTGTCTATAGCCCTTTCTT 3715  
DB 687 GCCTTCAAGACGTGATGCGCTGNNAGTTAATGATGCTTCTTCTCTGTGTCATAGGAT 628  
QY 3716 CCCACAGCATATAAAGCAGCACTCTGAAGCTTGGCAGAAAGGAGCTTATTATGATGCTG 3775  
DB 627 CNCCACAGCATGANAACGACTCCATTGAGGTNTGAGTTGAGCCAGGAGATGACCGCTGTG 568  
QY 3776 GTGAACCAATCCAGCAAGGTGGCGGTGACCCGCTTCCCTCTTTCCACCGTCCAGA 3835  
DB 567 GTGAGTTTCTCNCAGCCCTTGTCTTTGTGAGTTTGGTTTGGCCCTTACAGTCAAG 508  
QY 3836 ATGGTGTGTTAAAGCAGAACTCCCGTTTTCGCCAAGAGTCTAAGCAGCCCGTGAAGGA 3895  
DB 507 ACTGCATTCAAAAGAGCAGCTCTTGTGTTAGACCATCAATAGGTGTCCTGCTGCTGGA 448  
QY 3896 GCGGTGAACCCAGGTACAGTATGCTATCTCTTAAATAATTTTGGAGCTTGGAGGT 3955  
DB 447 TGCTCAAGCCCTTAAATGCTCTTGTGAGGCACTTGTACATGTTGCCAGACTTGGGGGT 388  
QY 3956 ATACTGTAGTCCGGAGCAACGCTAAAGGCTAGTCCGGTGGCTTGACC---GCGATGGTAA 4012  
DB 387 GCGTTCACTCTCTTTTGTACACTGAACATAAACCCTGTGCTTGACCAAGTCCATGGTAT 328  
QY 4013 GATCTCTGGCAAGGATCAGCACTTAAATATCTCCGTTTCGCAACATACGACCAC 4072  
DB 327 GGATCTTGGCAAGGATCAGCACTTAAATATCTCCGTTTCGCAACATACGACCAC 268  
QY 4073 ATAATTTCTATCTATCGGTGGATAAT 4098  
DB 267 AATACCTGATGTGAGGGTGGGTAGAT 242

Search completed: April 11, 2004, 07:57:51  
Job time : 13564 secs